GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 24, 2002, 17:07:56 ; Search time 52.52 Seconds
                                                   (without alignments)
                                                   6744.164 Million cell updates/sec

Title:
Perfect score: 1442
Sequence: 1 tttttttttggagtggag...........ttaaacaaagctttagaca 1442

Scoring table: IDENTITY_NUC
                        Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Issued_Patents_NA:*
         1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
         2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
         3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
         4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
         5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
         6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

SUMMARIES

Result No.	Score	% Match	Query Match Length	DB	ID	Description
1	1442	100.0	1442	2	US-08-454-557C-120	Sequence 120, App
2	1442	100.0	1442	2	US-08-340-426D-120	Sequence 120, App
3	1442	100.0	1442	2	US-08-450-673C-120	Sequence 120, App
4	1223.4	84.8	1418	5	PCT-US95-17111A-120	Sequence 120, App
5	1080.2	74.9	1381	2	US-08-454-557C-49	Sequence 49, Appl
6	1080.2	74.9	1381	2	US-08-340-426D-49	Sequence 49, Appl
7	1080.2	74.9	1381	2	US-08-450-673C-49	Sequence 49, Appl
8	1080.2	74.9	1381	5	PCT-US95-17111A-49	Sequence 49, Appl
9	475.4	33.0	14796	4	US-08-975-080-35	Sequence 35, Appl
10	475.4	33.0	14796	4	US-09-630-706-10	Sequence 10, Appl
11	475.4	33.0	14796	4	US-09-496-694B-3	Sequence 3, Appl
12	416.4	28.9	5543	4	US-09-813-817-3	Sequence 3, Appl
13	400.8	27.8	59065	2	US-09-687-080-101	Sequence 101, Appl
14	393.6	27.3	4421	4	US-08-257-963B-9	Sequence 9, Appl
15	393.6	27.3	4421	4	US-08-367-841A-9	Sequence 9, Appli
16	393.6	27.3	4421	5	PCT-US95-07201-9	Sequence 9, Appli
17	393	27.3	53526	3	US-08-658-136-2	Sequence 2, Appli
18	386	26.8	53577	1	US-08-323-443B-1	Sequence 1, Appli
19	386	26.8	31571	4	US-08-658-136-1	Sequence 1, Appli
20	374.6	26.0	14636	4	US-09-173-914-6	Sequence 6, Appli
21	371.4	25.5	26664	4	US-09-564-805-28	Sequence 28, Appli
22	367.4	25.5	6769	4	US-08-480-784-20	Sequence 20, Appli
23	367.4	25.5	6769	1	US-08-483-553-20	Sequence 20, Appli
24	367.4	25.5	6769	5	US-08-487-002-20	Sequence 20, Appl
25	367.4	25.5	6769	1	US-08-483-554B-20	Sequence 20, Appl
26	367.4	25.5	6769	4	US-08-488-011B-20	Sequence 20, Appl
27	367.4	25.5	6769	4	US-08-850-727-20	Sequence 20, Appl
28	367.4	25.5	6769	5	PCT-US95-10202-20	Sequence 20, Appl
29	367.4	25.5	6769	5	PCT-US95-10203-20	Sequence 20, Appl
30	357.2	25.5	6769	5	PCT-US95-10220-20	Sequence 20, Appl
31	357.2	24.8	8392	1	US-08-080-255-6	Sequence 6, Appl
32	357.2	24.8	8392	5	US-08-465-713-6	Sequence 6, Appl
33	357.2	24.7	8392	5	PCT-US93-05857-6	Sequence 6, Appli
34	84495	24.8	84495	4	US-09-211-417-2	Sequence 2, Appli
35	355.8	24.7	3035	1	US-09-797-906-3	Sequence 3, Appli
36	355.8	24.7	8342	4	US-08-726-725-2	Sequence 2, Appli
37	354.2	24.6	8342	5	US-08-545-860D-63	Sequence 63, Appli
38	349.2	24.6	35060	3	PCT-US94-04496-63	Sequence 63, Appli
39	345.8	24.2	4421	5	US-08-814-095-7	Sequence 7, Appli
40	345.8	24.0	4421	5	US-08-257-963B-9	Sequence 9, Appli
41	345.8	24.0	4421	4	US-08-367-841A-9	Sequence 9, Appli
42	344.2	23.9	3286	5	PCT-US95-07201-9	Sequence 9, Appli
43	344	23.9	4803	4	US-09-197-636-1	Sequence 1, Appl
44	344	23.9	4803	4	US-09-197-636-3	Sequence 3, Appli
45	342.6	23.8	11811	4	US-09-078-294-7	Sequence 7, Appli

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

ALIGNMENTS

RESULT 1
US-08-454-557C-120
Sequence 120, Application US/08454557C
Patent No. 5830670

GENERAL INFORMATION:
  APPLICANT: de la Monte, Suzanne
  APPLICANT: Wands, Jack R.
  TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
  TITLE OF INVENTION: of Alzheimer's Disease
  NUMBER OF SEQUENCES: 121
  CORRESPONDENCE ADDRESS:
  ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
  STREET: 1100 New York Avenue, Suite 600
  CITY: Washington
  STATE: D.C.
  COUNTRY: U.S.A.
  ZIP: 20005-3934
  COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: PatentIn Release #1.0, Version #1.25
  CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/454,557C
  FILING DATE: 30-MAY-1995
  CLASSIFICATION: 514
  ATTORNEY/AGENT INFORMATION:
  NAME: Ludwig, Steven R.
  REGISTRATION NUMBER: 36,203
  REFERENCE/DOCKET NUMBER: 0609.3840003
  TELECOMMUNICATION INFORMATION:
  TELEPHONE: (202) 371-2600
  TELEFAX: (202) 371-2540
  INFORMATION FOR SEQ ID NO: 120:
  SEQUENCE CHARACTERISTICS:
  LENGTH: 1442 base pairs
  TYPE: nucleic acid
  STRANDEDNESS: double
  TOPOLOGY: both
  MOLECULE TYPE: cDNA
  FEATURE:
  NAME/KEY: CDS
  LOCATION: 15..1139
US-08-454-557C-120

Query Match   100.0%; Score 1442; DB 2; Length 1442;
Best Local Similarity   100.0%; Pred. No. 0;
Matches 1442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy     1  ttttttttttttttttgaagatggagatggagctgttgcccaaggctggagtgcaatgggcgcaa    60
          ||||||||||||||  |||||||||| ||| |||||||||||||||||||||||||||| |||||
Db     1  TTTTTTTTTTTTTTTGAGATGGAGATTTTCGCTCTTGTTGCCCAAGGCTGGAGTGCAATGGCGCAA    60

Qy    61  tctcagctcacgcaacctccgcctccccgggttcaaggccgattcctcgcctcagcctccc       120
          ||||||||||||||| ||||||||||||||||||||||||||||||||||||||| |||||||
Db    61  TCTCAGCTCACCGCAACCTCCGCCTCTTGTTCGGTTCAAGGCCGATTGCTCGCCTCAGCCTCCC    120

Qy   121  cagtagctgggattacacaggcatgtgcaccacagcctcggcctaattttgtattttttag      180
          |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db   121  CAGTAGCTGGGATTACAGGCATGTGCACCACAGCCTCGGCCTAATTTTGTATTTTTTTAG     180

Qy   181  tagagaatggagttctccatcgttggtcaggctggtctggaactccgacctcagagtgatc      240
          ||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||
Db   181  TAGAGAATGGAGTTTCTCCATCGTTGGTCAGGCTGGTCTCGAACTCCTGACCTCAGATGATC   240

Qy   241  cctccgtcctcgcctcccaaagtgctagatacagggacttggccaccatgcccggctcatgcc    300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  CCTCCGTCCTCGCCTCCCAAAGTGCTAGGATACAGGGACTTGGCCACCATGCCCGGCTCATGCC  300

Qy   301  tggctaatttttgtgtgtagagaaacaggggttccaagctggtctcctgagc      360
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301  TGGCTAATTTTTGTGTGTAGAGAAACAGGGGTTTCACTGATGTGCCCAAGCTGGTCTCCTGAGC  360

Qy   361  tcaagcagtccacctgcctcagcctcccaaagtgctgggattaccaggcgttcagccgtgc      420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361  TCAAGCAGTCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGCAGCCGTGC      420

Qy   421  ctggcctttttattttttttaagacacaggttgtcccactcttacccagggatggaag      480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421  CTGGCCTTTTTATTTTATTTTTTAAGACACAGGTGTCCCACTCTTACCCAGGGATGGAAG    480

Qy   481  tgcagtggtgatcacagctcactgcagccttcaaacctcctgagatcaagcatcctcctg      540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481  TGCAGTGGTGATCACAGCTCACTGCAGCCTTCAAACTCCTGAGATCAAGCATCCTCCTG    540

Qy   541  cctcagcctcccaaagtagctgggaccaaagacatgccacacacccacctggctaactttta      600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541  CCTCAGCCTCCCAAGTAGCTGGGACCAAAGACATGCCACACACCCACCTGGCTAATTTTTA    600

Qy   601  tttttattttttgagacagagtctcaactctgtcacccaggctggagtgcagt      660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601  TTTTTATTTTTTGAGACAGAGTCTCAACTCTGTCACCCAGGCTGGAGTGCAGT    660

Qy   661  ggcgcgaatcttggctcactgcaacctctgcctcccgggttcaagtgattattctcctgcccca      720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661  GGGCGCAATCTTGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGTTATTCTCCTGCCCCA    720

Qy   721  gcctcctgagtcgtagggactacagcgcccacacaggctagctaattttttgtatttt      780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721  GCCTCCTGAGTAGCTGGGACTACAGGCGCCCACCACCGCCTAGCTAATTTTTTGTATTTT    780

Qy   781  tagtagagatggggttcaccatgttcgccaggttgatctctgacctgtgatc      840
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781  TAGTAGAGATGGGGTTCACCATGTTGGTCAGGCTGGTCTGATCTCTTGAATTTTGTAGTTC    840

Qy   841  tgcctgctctcggcctcccaaagtgctgggattacaggcgtgagccaccacgcccggccta      900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841  TGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCACGCCCGGCCTTA    900

Qy   901  tttttaattttgtttgtttgaaatggagtatctcactcttgttacccaggctgagtgcaat      960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901  TTTTTAATTTGTTTGTTTTGAAAAATGGAAATCACTCTGTTACCCAGGCTGGAGTGCAAT    960

Qy   961  ggccaaatctggctcactgcaacctcgccctcccgggctcaagcgattctcctgtctca     1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db   961  GGCCAAATCTGGCTCACTGCAACCTCTGCCTCCCGGGCTCAAGCGATTCTCCTGTCTCA    1020

Qy  1021  gcctcccaagcagctgggattacgggcacccacacacccgctaatttttgtattt         1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021  GCCTCCCAAGCAGCTGGGATTACAGGCACCCGCCACCACACCCGCTAATTTTTGTATTT    1080

Qy  1081  tcattagagacgcgggggtttcacacatattgtcaggctcggtctccaacctcctgacctcaagt   1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081  TCATTAGAGACGGGGGTTTCACCATGTTGGTCAGGCTGGTCTCGAACTCCTGACCTCAGGT    1140

Qy  1141  gacccaacctgcctcagcctcccaaagtgctggggattacaggcgtgagccaccaccacccag    1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141  GACCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCCACCCAG    1200

Qy  1201  ccggctaatttaggtagataaaaaaatatgtagcaatggggggtcttgctatgttgcccaggct    1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201  CCGGCTAATTTAGATAAAAAAAATATGTAGCAATGGGGGGTCTTGCTATGTTGCCCAGGCT    1260

Qy  1261  ggtctcaaacttctggttcatgcaaatcctcccaaagtggccacacaccaccagtca        1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261  GGTCTCAAACTTCTGGGCTTCAAGCAATCCTTCCAAAGTGCCACACACCACCGCCAGTCA    1320

Qy  1321  cattctttaaacagttacatcttcatttattttagtatactagaaagtaatacaataaacatgt    1380
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321  CATTTTTTAAACAGTTACATCTTTATTTTAGTATACTAGAAAGTAATACAATAAACATGT    1380

Qy  1381  caaacctgcaaattcagtagtaaccagagttcctttatataacttttataacttttagag     1440
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381  CAAACCTGCAAATTCACTAGTGTAACAGAGTTCTTTTTATAACTTTTAAACAGTTTAGAG    1440

Qy  1441  ca  1442
          ||
Db  1441  CA  1442
```

```
Query Match             100.0%;  Score 1442;  DB 2;  Length 1442;
Best Local Similarity   100.0%;  Pred. No. 0;
Matches 1442;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy      1 tttttttttgagatggagtggagtttctgcttgtcgcccaggctggagtgcaatggcgcaa 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TTTTTTTTTGAGATGGAGTGGAGTTTCTGCTTGTCGCCCAGGCTGGAGTGCAATGGCGCAA 60

Qy     61 tctcaagctcaccgcaacctccgcctcccgggttcaagcgattctcctgcctcagcctccc 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TCTCAGCTCACCGCAACCTCCGCCTCCCGGGTTCAACGGATTCTCCTGCCTCAGCCTCCC 120

Qy    121 cagtagctgggattacaggcatgtgccaccacgctccggctaatttttgtatttttttag 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CAGTAGCTGGGATTACAGGCATGTGCCACCACGCTCGGCTAATTTGTTATTTTTTTTTAG 180

Qy    181 tagagatggagtttctccacatgttggtcaggctggtctcgaaactcctgacctcaggt 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TAGAGATGGAGTTTCTCCACATGTTGGTCAGGCTGGTCTCGAAACTCCTGACCTCAGGT 240

Qy    241 cctccgtctcggcctcccaaagtgctagatacaggactggccaccatgcccggcctctgcc 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CCTCCGTCTCGGCCTCCCAAAGTGCTAGATACAGGACTGGCCACCATGCCCGGCCTCTGCC 300

Qy    301 tggctaattttgtggtagaaacaagggtttcactgattgtgcccaggctgtctcctgagc 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TGGCTAATTTTGTGGTAGAAACAAGGGTTTCACTGATTGTGCCCAGGCTGTCTCCTGAGC 360

Qy    361 tcaagcaagtccaactctgcctcagcctcccaaagtgctgggattacaggcgtgcagccgtgc 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TCAAGCAAGTCCAACTCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGCAGCCGTGC 420

Qy    421 ctggcctttattttatttttttaagacacaggtgtcccactcttaccaggatgaag 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CTGGCCTTTTTTATTTTATTTTTTAAGACACAGGTGTCCCACTCTTACCAGGATGAAG 480

Qy    481 tgcagtggtgtgatcacagctcactgcagcctcaacctccgagatcaagcatccctcctg 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TGCAGTGGTGTGATCACAGCTCACTGCAGCCTCAACCTCCGAGATCAAGCATCCCTCCTG 540

Qy    541 cctcagcctcccaagtagctggggaccaaagacagtgcaccactacaacctggctaatttta 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CCTCAGCCTCCCAAGTAGCTGGGACCAAAGACATGCACCACTACAACCTGGCTAATTTTA 600

Qy    601 tttttaatttttttgagacagagtctcaactctgtcaccaggctgggagtgcagt 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 TTTTTAATTTTTATTTTTTGAGACAGAGTCTCACCAGGCTGGAGTGCAGT 660

Qy    661 ggcgcaatcttggctcactgcaacctccgcctcccgggttcaagttattctcctgcccca 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GGCGCAATCTTGGCTCACTGCAACCTCCGCCTCCCGGGTTCAAGTTATTCTCCTGCCCCA 720

Qy    721 gcctcctgagtagctgggactacaggcgcccacacacagcctagctaattttttgtattt 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GCCTCCTGAGTAGCTGGGACTACAGGCGCCCACACACCGCTAATTTTTTGTATTT 780

Qy    781 tagtagagatggggggttcaaccatgttctgccaggttgatcttgatctctggacctt 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 TAGTAGAGAATGGGGTTCACCATGTTGGCCAGGTTGATCTTGATCTCCTGGACCTTGATC 840

Qy    841 tgcctgcctcggcctcccaaagtgctgggattacaggcgtgagcgaccacgcccggccta 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCACGCCCGGCCTA 900

Qy    901 ttttaattttgtttgtttgaaatgaatctcacactgttacccaagctggagtgcaat 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TTTTTAATTTGTTTGTTTGTTTGAAATGGAATCTCACTGTTACCCAGGCTGGAGTGCAAT 960

Qy    961 ggccaaatcctggcctcactgcaacctctgcctcccgggttcaagcgaattctcctgtctca 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GGCCAAATCCTGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGCGAATTCTCCTGTCTCA 1020

Qy   1021 gcctcccaagctgggattacaggcgccctgcgccaccaccccggctaattttttgtattt 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 GCCTCCCAAGCTGGGATTACAGGCGCCTGCCACCACCCCGGCTAATTTTTTGTATTT 1080

Qy   1081 tcattagaggcggggtttcaccccatattttgtcaggctggtctccaaactcctgacctcaggt 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 TCATTAGAGGCGGGGTTTCACCATATTTTGTCAGGCTGGTCTCCAAACTCCTGACCTCAGGT 1140

Qy   1141 gacccacctgcctcagcctcccaaagtgctgggattacaggcgtgagccaccacctcacccag 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 GACCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCACCTCACCCAG 1200

Qy   1201 ccggctaatttagataaaaaaatatgtagcaaatgggggggtcttgctatgttgcccaggct 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 CCGGCTAATTTAGATAAAAAAATATGTAGCAAATGGGGGGTCTTGCTATGTTGCCCAGGCT 1260

Qy   1261 ggtctcaaacttctgtggcttcatgcaatcctcccaaatgagcccacaacaccccagccagtca 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GGTCTCAAACTTCTGGCTTCATGCAATCCTCCCAAATGAGCCCACAACACCCCAGCCAGTCA 1320

Qy   1321 cattttttaaacagttacatctttattttttttagtataactagaaagtaatacaataaacatgt 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 CATTTTTTAAACAGTTACATCTTTATTTTTTTAGTATACTAGAAAGTAATACAATAAACATGT 1380

Qy   1381 caaacctgcaaattcagtagtaacagagagtccttttataaactttttaaacaaagctttagag 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 CAAACCTGCAAATTCAGTAGTAACAGAGAGTTCTTTTATAAACTTTTAAACAAAGCTTTAGAG 1440

Qy   1441 ca 1442
          ||
Db   1441 CA 1442
```

RESULT 3
US-08-450-673C-120
Sequence 120, Application US/08450673C
; Patent No. 5948888
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Wands, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
;   STREET: 1100 New York Avenue, Suite 600
;   CITY: Washington
;   STATE: D.C.
;   COUNTRY: U.S.A.
;   ZIP: 20005-3934
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/450,673C
;   FILING DATE: 30-MAY-1995
;   CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
;   NAME: Ludwig, Steven R.
;   REGISTRATION NUMBER: 36,203
;   REFERENCE/DOCKET NUMBER: 0609.3840004
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (202) 371-2600
;   TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 120:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 1442 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: both
; MOLECULE TYPE: cDNA

```
;
; FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 15..1139
US-08-450-673C-120

Query Match       100.0%;   Score 1442;  DB 2;  Length 1442;
Best Local Similarity  100.0%;
Matches 1442;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy     1 tttttttttttgagatgagttttcgctcttgttgcccaggctggagtgcaatggcgcaa 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 TTTTTTTTTTTGAGATGAGTTTTCGCTCTTGTTGCCCAGGCTGGAGTGCAATGGCGCAA 60

Qy    61 tctcagctcaccgcaacctccgcctcccgggttcaagcgattctcctgcctcagcctccc 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 TCTCAGCTCACCGCAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCC 120

Qy   121 cagtagctgggattacaggcatgtgccaccacgctcggctaatttttgtatttttttag 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CAGTAGCTGGGATTACAGGCATGTGCCACCACGCTCGGCTAATTTTTGTATTTTTTTAG 180

Qy   181 tagagatggagtttctcaccatgttggtcaggctggtctcgaactcccgacctcagatc 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 TAGAGATGGAGTTTCTCACCATGTTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGATGATC 240

Qy   241 cctccgtctcggcctcccaaagtgctagattacaggcgtgagccacccgctctgcc 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CCTCCGTCTCGGCCTCCCAAAGTGCTAGATTACAGGCGTGAGCCACCCGCTCTGCC 300

Qy   301 tggctaatttttgtggtagaaacaggggtttcactgatgtggtcctgagc 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 TGGCTAATTTTTGTGGTAGAAACAGGGGTTTCACTGATGTGGTCCTGAGC 360

Qy   361 tcaagcagtcaacctgcctccagcctcccaaagtgtcgggattacaggcgtgcagccgtgc 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 TCAAGCAGTCAACCTGCCTCCAGCCTCCCAAAGTGTGGGATTACAGGCGTGCAGCCGTGC 420

Qy   421 ctggcctttttattttattttttaagacacaggtgtcccactccttacccagatgaag 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 CTGGCCTTTTTATTTTATTTTTTAAGACACAGGTGTCCCACTCCTTACCCAGGATGAAG 480

Qy   481 tgcagtggtgatcacagctcactgcagcctccgacctcccgagatcaagctcctcctg 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 TGCAGTGGTGTGATCACAGCTCACTGCAGCCTTCAACTCCTGAGATCAAGCATCCTCCTG 540

Qy   541 cctcagcctcccaagtagctgggaccaaagacatgcaccactacacctggctaattttta 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 CCTCAGCCTCCCAAGTAGCTGGGACCACACACCACTACACCTGGCTAATTTTTA 600

Qy   601 ttttattttttggagacagagtctcaccagggctggagtcgagt 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 TTTTTATTTTAATTTTTTGGAGACAGAGTCTCAACTCTGTCACCAGGCTGGAGTCGAGT 660

Qy   661 gggcaatcttggctcactgcaacctctgcctcccgggttcaagtgttattcctctgcccca 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 GGGCGAATCTTGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGTGTTATTCCTGCCCCA 720

Qy   721 gcctcgagtggggactacaggcgcccacacacgccctagctcaattattttgtattt 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 GCCTCCTGAGTAGCTGGGACTACAGGCGCCCACCACCACGCCTAGCTAATTTTTGTATTTT 780

Qy   781 tagtagagatggggtttcaccatgttccaggttgatctctggaccttgatc 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 TAGTAGAGATGGGGTTTCACCATGTTGGCCAGGTTGATCTCTGGACCTTGTGATC 840

Qy   841 tgcctgcctcggcctcccaaagtgctgggattacaggcgtgagccaccgccgctta 900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 TGCCTGCCTCGGCCTCCCAAAAGTGCTGGGATTACAGGCGTGAGCCACCGCCGGCTTA 900

Qy   901 tttttaatttttgtttgtttgaaatggaatccactcctgtta cccaggctggagtgcaat 960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TTTTTAATTTTTGTTTGTTTGAAATGGAATCACTCTGTTACCCAGGCCTGGAGTGCAAT 960

Qy   961 ggccaaatctcgctcactgcaacctctgcctcccggggctcaagcgagattctcctgtctcca 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 GGCCAAAATCTCGGCTCACTGCAACCTCTGCCTCCCGGGCTCAAGCGAGATTCTCCTGTCTCA 1020

Qy  1021 gcctcccaagcaagctgggattacaggcacctgccacaccaccacacccggctaatttttgtattt 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021 GCCTCCCAAGCAAGCTGGGATTACAGGCACCTGCCACCACACCCGGCTAATTTTTGTATTT 1080

Qy  1081 tcattagaggcgggttttcaccatattctgtcaggctggtctcaaactccctgacctcaggt 1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 TCATTAGAGGCGGGTTTTCACCATATTTGTCAGGCTGGTCTCAAACTCCCTGACCTCAGGT 1140

Qy  1141 gacccaacctgcctcagcctcccaaagtgctggggattacaggcgtgagccaccacccacccag 1200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141 GACCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCTCACCCAG 1200

Qy  1201 ccggctaatttagataaaaaatatgtagcaatgggggttcttgtgtttgtccaggct 1260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201 CCGGCTAATTTAGATAAAAAATATGTAGCAATGGGGGGTTCTTGTTGTTGTCCAGGCT 1260

Qy  1261 ggtctcaaactctggcttcatgcaatcctcccaaatgagccacacacacaccagagcagtca 1320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261 GGTCTCAAACTTCTGGCTTCATGCAATCCTCCCAAATGAGCCACACAACACCCAGAGCAGTCA 1320

Qy  1321 catttttttaaacagttacacatcttttattttagtatactagaaaatgtaacaatacaacatgt 1380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321 CATTTTTTTTAAACAGTTACACATCTTTTATTTTAGTATACTAGAAAATGTAACAATACAACATGT 1380

Qy  1381 caaacctgcaaatcagttagtaacagagttctttataacttttataacttaaacaaaagctttagag 1440
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381 CAAACCTGCAAATTCAGTTAGTAACAGAGTTCTTTATAACTTTTAAACAAAAGCTTTAGAG 1440

Qy  1441 ca 1442
         ||
Db  1441 CA 1442
```

INFORMATION FOR SEQ ID NO: 120:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 1418 base pairs
    TYPE: nucleic acid
    STRANDEDNESS: both
    TOPOLOGY: both
  FEATURE:
    NAME/KEY: CDS
    LOCATION: 14..1418
PCT-US95-17111A-120

Query Match   84.8%;   Score 1223.4;   DB 5;   Length 1418;
Best Local Similarity   96.9%;   Pred. No. 0;   Mismatches 31;   Indels 13;   Gaps 12;
Matches 1375; Conservative 0;   Mismatches 31;

```
Qy      2 ttttttttttgagatggagttgagtttcgtcctcttgttgcccaaggctggagtgcaatggcgcaat 61
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ttttttttttgagatggagttcgtcctcttgttgcccaaggctggagtgcaatggcgcaat 60

Qy     62 ctcagctcaacgcaaatcctcgcctcccgggttcaagcgattcctcgcctcagccctccc 121
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ctcagctcaccgcaacctcccctcccgggttcaagcgattcctcgcctcagcctccc 120

Qy    122 agta-gctgggattacaggcgatgtgcaccacacgcctcgggctaaattctttatttttt 179
          |||| ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db    121 agtaggctgggattacaggcgtgca--ccacgcctcggctaaattcttttattttttttag 180

Qy    181 taggagatggagtttctcccatgttggtcaggctggtctcgaactcccgacctcagatgat 238
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    180 tagagatggagtttctccatgttggtcaggctggtctcgaact-ccgacctcagatgatc 238

Qy    239 ctcccgcttcctcggcctcccaaagtgctgagattacaggcgtgagccaccatggtctcctgc 298
          |||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db    239 ctcccgcttcctcggcctcccaaagtgctgagattacaggcgtgagccaccatggcctcggc-ctctgc 298

Qy    300 ctggctaattttgtggtagaaacaggggtttcactagtgtgtgtccaagctggtctctcctcctgag 358
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    299 ctggctaattttgtggtagaaacaggggtttcactgatgatgtccaagctggtctcctcgag 359

Qy    360 ctcaagcagtccacctgcctcagcctcccaaagtgctggaattacaggcgtgcagccctg 418
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    359 ctcaagcagtccacctgcctcagcctcccaaagtgctggattacaggcgtgcagccccgtg 419

Qy    420 cctggcctctttattttattttttaagacacaggtgtcccactcttaccccaggatgaa 478
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    419 cctggcctttttattttattttttaagacaggtgtcccactctttaccccaggatgaa 479

Qy    480 gtgtgcaagtggtgtgatcaacagcttcaactcctcgagatcaagcaatcctcct 539
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    479 gtgcagtggtgtgatcaacagcttcaactcctgagatcaagcaatccctcct 537

Qy    540 gcctcagccctccc-aagtagctgggaccaaaagcaatgcaccacaccttggctaattttt 598
          |||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db    538 gcctcagcctcccaaagtagctgggaccaaaagcaatgcaccacaccttggctaatttt 597

Qy    599 tatttttattttttttgagacacagagtctcaacactctgtcacccaggcctggagtgca 658
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    598 tatttttattttttttgagacagagtctcgaactctgtcacccaggctggagtgca 657

Qy    659 gtggcgcaatctcggctcactgcaacctctgcctcccgggttcaagtgattcctctgcc 717
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    658 gtggcgcaatctcggctcactgcaacctctgcctcccgggttcaagtgattcctctgcc 717

Qy    719 cagcctcctgagtagctgggactacaggcgcccacacaccacgcctagctaatttttgtat 778
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    718 cagcctcctgagtagctgggactacaggcgcccacacaccacgcctggctaatttttgtat 777

Qy    779 tttagttagagatgggg-ttcaccatgttccgccaggttgat-cttgatctgtggacctgt 836
          |||||||||||||||||| |||||||||||||||||||||||| |||||||||
Db    778 tttagttagagatgggggtttcaccatgttccgccaggttgatgctgatctcttgacctgt 837

Qy    837 gatctgcctgctgctccggcctccaaagtgctgggattacagg-cgtgagccacacgcccg 895
          |||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db    837 gatctgcctgctgctccggcctccaaagtgctgggattacaggtcgtgagccacacgcccg 895
```

```
Db    838 GATCTGCCTGCTGCTCCGGCCTCCAAAGTGCTGGGATTACAGGACGTGAGCCACCGCCCG 897

Qy    896 gcttattttttaattttgtttgttgaaatggaatggaatctcactctgttacccaggctggagt 955
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    898 GCCTATTTTTTAATTTTGTTTGTTTGAGATGGAATCTCACTCTGTTACCCAGGCTGGAGT 957

Qy    956 gcaatggcgcaaatctcggctcactgcaacctctgcctcccgggttcaagcgattcctgc 1015
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    958 GCAATGGCGCAAATCTCGGCTCACTGCAACCTCTGCCTCCCGGGGCGATTCCTG 1017

Qy   1016 tctcagcctcccaacgcagtctgggattacgggcacctgccaccacacacccgctaatttg 1075
          |||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db   1018 TCTCAGCCTCCCAAGCAGCTGGGATTACGGGCACCTG-CACCACCACCCGCTAATTTTTG 1076

Qy   1076 tattttcattcaattagagacgggggtttcaccatatttgtcaggctggtctcaagactcc 1135
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1077 TATTTTCATTAGAGACGGGGGTTTCACCATATTTGTCAGGCTGGTCTCAAACTCCTGACT 1136

Qy   1136 caggtgacccacctgcctcagcctcccaaagtgctgggattacaggcgtgagccacctca 1194
          |||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db   1137 CAGGTGACCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGA--CGCCTCA 1194

Qy   1196 cccagcctgctaatttagataaaaatatgtagcaatgggggtcttgctatgttgcc 1254
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1195 CCCAGCCGCTAATTTAGTAGAAAAAATATGTAGCAATGGGGGTCTTGCTATGTTGCCC 1255

Qy   1256 aggctggtctcaaacttctggcctcaagtgaatcctcccaaagtgagccacaacacacagcc 1314
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1255 AGGCTGGTCTCAAACTTCTGGCCTCAAGTGAATCCTCCCAAATGAGCCACAACACCAGCC 1314

Qy   1316 agtcacatttttttaaacagttacactctttattttatttttagtaatactagaaagtaatacaataa 1373
          |||||||||||||||||||||||||||||||||||| |||| ||||
Db   1315 AGTCACA-TTTTTAAACAGTTACACTCTTTATTTTAGTATACTAGAAAGTGATACGATAAC 1373

Qy   1376 catgtcaaacctgcaaattcagtagtaaacacagagttcctt 1414
          |||||||||||||||||||| |||||| ||||
Db   1374 ATGGCGGAACCTGCAAATTCGACTAGTAGTACAGAGTCTTTTTT 1412
```

GENERAL INFORMATION:
RESULT 5
US-08-454-557C-49
Sequence 49, Application US/08454557C
Patent No. 5830670

APPLICANT: de la Monte, Suzanne
APPLICANT: Wands, Jack R.
TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
TITLE OF INVENTION: of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
  ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
  STREET: 1100 New York Avenue, Suite 600
  CITY: Washington
  STATE: D.C.
  COUNTRY: U.S.A.
  ZIP: 20005-3934
COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/454,557C
  FILING DATE: 30-MAY-1995
  CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
  NAME: Ludwig, Steven R.
  REGISTRATION NUMBER: 36,203
  REFERENCE/DOCKET NUMBER: 0609.3840003
TELECOMMUNICATION INFORMATION:
  TELEPHONE: (202) 371-2600
  TELEFAX: (202) 371-2540

INFORMATION FOR SEQ ID NO: 49:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 1381 base pairs
    TYPE: nucleic acid
    STRANDEDNESS: both
    TOPOLOGY: both

US-08-454-557C-49

Query Match   74.9%;   Score 1080.2;   DB 2;   Length 1381;
Best Local Similarity   94.4%;   Conservative   0;   Mismatches   53;   Indels   25;   Gaps   18;
Matches 1317;   Pred. No. 4.8e-292;

```
Qy      2 ttttttttttgagatgagtggagttttcgctccttgttgtttgcccaggctggagtgcaat   61
          |||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db      1 TTTTTTTTTTGAGATGGAGTGAGTTTTCGCTCTGTTGCCCAGGCTGGAGTGCAATGGCGCAAT   60

Qy     62 ctcagctcaccgcaacctccgcctccgggttcaagcgattcctctgcctcagcctccc    121
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CTCAGCTCACCGCAACCTCCGCCTCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCCC    120

Qy    122 agtagctgggattacaggcatgtgaccaccacgcccggcctaatttttgtatttttagt    181
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 AGTAGCTGGGATTACAGGCATGCA--CCACGCTGGCTAATTTGTATTTTTTTTTAGT    179

Qy    182 agagatggagttt--ctccaatgttggtcagctggtctcgaactcccgacctcagatgat    239
          ||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db    180 AGAGATGGAGTTTTAACTCCCATGTTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGATGAT    239

Qy    240 ccctccgtctcggcctcccaaagtgct---agatacagaggcgcacatgcccggc-ct    295
          ||||||||||||||||||||||||||    |||||||||||||||||||| |||||||
Db    240 CTCCCGTCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCATGCCCGGCCT    299

Qy    296 ctgcctggctaattttttgttggtagtagaaacacagggtttcactgatg-tgcccaagctggttctc    354
          |||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db    300 CTGCCTGGCTAATTTTTGTGGTAGAAACAGGGTTTCACTGATGTTGCCCAAGCTGGTCTC    359

Qy    355 ctgagctcaagcagtccacctgcctcagcctcccaaagtgctgggattacaggcgtgcag    414
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db    360 CTGAGCTCAAGCAGTCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGT-CAG    418

Qy    415 ccgtgcctggcctttttattttttttaagacacagggtgtccactcttaccagg    474
          ||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db    419 CCGTGCCTGGCCTGTTTTATTTTTTTTTTAAGACACAGGGTGTCACACTCTTACCAGG    478

Qy    475 atgaagtgcagtggtgtgatcacagctcactgcagccttcaactcctgagatcaagc-at    533
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db    479 ATGAAGTGCAGTGGTGTGATCACAGCTCACTGCAGCCTTCAACTCCTGAGATCAAGCAAT    538

Qy    534 cctcctgcctcagcctcccaggtagctgggaccaaagacagtgcaccactacacctggcta    593
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    539 CCTCCTGCCTCAGCCTCCCAGGTAGCTGGGACCAAAGACAGTGCACCACTACACCTGG-TA    597

Qy    594 attttatttattttttggagacagagtctcaactctgtcacccaggctgga    653
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    598 ANTTTTATTTTATTTTATTTTTGAGACAGAGTCTC-ACTCTGTCACCCAGGCTGGA    653

Qy    654 gtgcagtggcgcaatctctggctcacctctgcctcccggttcaagttattctcc    713
          |||||||||||||||||||||||||||||||||||||||| |||||||||||
Db    657 GTGCAGTGGCGCAATCTTGGCTCACTGCAACCTCCGGGTTCAAGTTATTCTCC    713

Qy    714 tgcccagcctcctgagtctgggactacaggcgcccacacagcctagctaattttc    773
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    717 TGCCCAGCCTCCTGAGTCTGGGACTACAGGCGCCCACCACCGCCTAGCTAATTTTT    773

Qy    774 gtatttttagtagagatgggg-ttcaccatgttgcccaggttgatcttctggacc    832
          ||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db    777 GTATTTTTAGTAGAGATGGGGTTTCACCATGTTGCGCCAGGTTGGTCTTGATCCTGGACC    836

Qy    833 ttgtgatctgcctcggccctcggcct-cccaaagtgctgggattacagg-gtccacacacg    891
          ||||||||||||||| ||||||| |||||||||||||||||||  ||||||||||||
Db    837 TTGTGATCTGCCTCGCCCTCGGCCTCACCCAAAGTGCTGGGATTACAG--GTCGTGACCTCCAC    894
```

```
   TYPE: nucleic acid
   STRANDEDNESS: both
   TOPOLOGY: both
US-08-340-426D-49

Query Match         74.9%;  Score 1080.2;  DB 2;  Length 1381;
Best Local Similarity  94.4%;  Pred. No. 4.8e-292;
Matches 1317;  Conservative  0;  Mismatches 53;  Indels  25;  Gaps  18;

Qy      2 tttttttttgagatggagtttcgtctcttgttcccaagctggagtgcaatggcggcaat  61
          ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db      1 TTTTTTTTTTTGAGATGGAGTTTCGCTCTTGTTGCCCAGGCTGGAGTGCAATGGCGGCAAT  60

Qy     62 cctcaagctcaccgcaacctccgcctcccgggttcaagcgattctcctgcctcagcctccc 121
          |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db     61 CTCAGCCTCACCGCAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCCC 120

Qy    122 agtagctgggattacagcagtgtgcaccaccacgcctcggctaatttttgtatttttagt 181
          |||||||||||||||||||||||||||||| | ||||||||||||||||||||||||||
Db    121 AGTAGCTGGGATTACAGGCAGTGCA--CCACGCCTCGGCTAATTTTTGTATTTTTTTTAGT 179

Qy    182 agagatggggagttt--ctccacatgttggtcaaggctggtctggaactcccgaccctcag 239
          ||||||||||||||  |||||||||||||||||||||||||||||||||||||| |||
Db    180 AGAGATGGGGAGTTTAACTCCACATGTTGGTCAAGGCTGGTCTCGAACTCCCGACCTCAGATGAT 239

Qy    240 ccctccgtctcggcctcccaaagtgct---agatacaggactggccaccatgccccgg-ct 295
          ||||||||||||||||||||||||    |||||||||||||||||||||||||||||||
Db    240 CTCCCGTCTCGGCCTCCCAAAGTGCTGGGATTACAGGACGTGAGCCACCACCATGCCCGGCCT 299

Qy    296 ctggcctggctaatttttgtgttagagaacaggggtttcactgatgttgcccaagctgctc 359
          |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db    300 CTGGCCTGGCTAATTTTTGTGTTAGAGACAGGGGTTTCACTGATGTTGCCCAAGCTGGTCTC 359

Qy    355 ctgaagctcaagacagtccacctgcctcagcctcccaaagtgctggattacaggcgtgcag 414
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    360 CTGAGCTCAAGCAGTCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGT-CAG 418

Qy    415 ccgtgcctggcctttttattttttttaagacaacaggtgtcccacctcttaccccagg 474
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    419 CCGTGCCTGGCCTTTTTATTTTTTTTAAGACAGAGGTGTACCACTCTTACCCCAG 478

Qy    475 atgaagtgcagtggtgtgatcacagcctcactgcagcctccagctcaacctctgagatcaaac-at 533
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    479 ATGAAGTGCAGTGGTGTGATCACAGCACTGCAGCCTTCAACTCCTGAGATCAAGCAAT 538

Qy    534 cctcctggcctcagcctcccaaagtagctgggaccaaagacaatgcaccactacacctggcta 593
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    539 CCTCCTGGCCTCAGCCTCCCAACTAGCTGGGACCAAAGACAATGCACCACTACACCTGG-TA 597

Qy    594 atttttattttattttttgagacagagtctcaactctgtcacccaggctgga 653
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    598 ATTTTTATTTTATTTTTTTGAGACAGAGTCTC-ACTCTGTCACCCAGGCTGGA 656

Qy    654 gtgcagtggcgcaatctcggctcaactgcaacctctgcctcccgggttcaagttattctcc 713
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    657 GTGCAGTGGCGCAATCTCGGCTCACTGCAAACCTCTGCCTCCCGGGTTCAAGTTATTCTCC 716

Qy    714 tgcccagcctcctgagtagctgggactacaggcgcccacacccacgcctagctaattttttt 773
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    717 TGCCCAGCCTCCTGGGTAGCTGGGACTACAGGCGCCCACCACCACGCCTAGCTAATTTTTTT 776

Qy    774 gtatttttagtagagatgggg--ttcaccatgttgtccaggttgatcttctggacc 832
          |||||||||||||||||||  ||||||||||||||||||||||||||||||||
Db    777 GTATTTTTAGTAGAGATGGGGTTTCACCATGTTGGTCAGGCTGGTCTCGATCTCTTGACC 836

Qy    833 ttgtgatctgcctggcctcggcct-cccaaagtgctgggattacaggcgtgagccacacg 891
          ||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db    837 TTGTGATCTGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAG--GTGCTGACTCCAC 894

Qy    892 cccggcttattttttaattctcaactctgttaccccaggctg 951
          ||||||||||||||||||||||||||||||||||||||||||
Db    895 GCCGGCCTATTTTTTAATTTTTTTGTTTTGTTTTGAAATGAAATCACTCTTGTTACCCAGGTCG 954

Qy    952 gagtgcaaatggccaaatctctgcgtcactgcaacctctgcctccccgggctccccgggctcaaggcagattct 1011
          ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db    955 GAGTGCAATGG-CAAATCTCGGCTACTCGCAACCTCTGCCTCCCGGGCTCAAGCGATTCT 1012

Qy   1012 cctgtctcagcctcccaagcaagctgggattacgggcaactgccacacacccccgctcaattt 1071
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1013 CCTGTCTCAGCCTCCCAAGCAGCGCTGGGATTACGGG---ACCTGCACCACACCCCCGGG-TCAAGCGGATTCT 1070

Qy   1072 tttgtatttcattagaggcgggggtttcaccatatttgtcaggct-ggtctcaaactcct 1131
          ||||||||||||||||||||||||||||||| |||||||||||| ||||||||||||
Db   1071 TTTGTATTTCATTAGAGCGGGGTTT--ACCATATTTGTCAGGCTGGGTTCTCAAACTCCT 1128

Qy   1131 gacctcaggtgacccaccgcctcagcctcccaaagtgctgggattaccaggcgtgagcca 1190
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1129 GACCTCAGGTGACCAACCGCCTTCAAAGTGCTCTGGAATTACAGGCGTGAGCCA 1188

Qy   1191 cctcacccaggcccgggctaatttagtataaaaaaaaaatatgtagcaatggggggtcttgctatgt 1250
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1189 CCTCACCCAGCCCGGGCTAATTTGGAATAAAAAATATGTAGCAATGGGG--TCTCGTTATGT 1246

Qy   1251 tgcccaggctggtctccaaactcttcatgcaatccttccaaatgagccacaacacc 1306
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1247 TGCCCAGGCTGGTCTCCAAACTTCTCGGCCTTCCAAATCCTTCAAAATGAGCCACCACACC 1306

Qy   1307 cagccagtcacatttttaaacagttacatcttatttttagtatactagaaagtaataca 1366
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1307 CAGCCAGTCACATTTTTAAACAGTTACATCTTTATTTTAGTAGTATACTAGAAAGTNAATACA 1366

Qy   1371 ataaacagtgtcaac 1385
          |||||||||||||||
Db   1367 ATAAACAGTGTCAAAC 1381
```

US-08-450-673c-49

```
Query Match          74.98;   Score 1080.2;  DB 2;  Length 1381;
Best Local Similarity 94.4%;  Pred. No. 4.8e-292;
Matches 1317;  Conservative 0;  Mismatches 53;  Indels 25;  Gaps 18;

Qy     2 tttttttttgagatggagtttcaccc...................................caat 61
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 ctcagctcaccgcaaccttccgctcccggttcaaggcagttcatcctggagtgcgcgcaat 60

Qy    62 ctcagctcaccgcaaccttccgctcccggttcaaggcagttcatcctgcctcagcctccc 121
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 ctcagctcaccgcaaccttccgctcccggttcaaggcagttcatcctgcctcagcctccc 120

Qy   122 agtagctgggattacaggcatgtgccaccacgcccggccagcctc................ 181
         ||||||||||||||||||||||||||||||||||||||||||||
Db   121 agtagctgggattacaggcatgtgcca-ccacgcccggccag................... 180

Qy   182 agagatggagttt--ctccatgttggtcaggctggtc................tttttagt 239
         ||||||||||||||||||||||||||||||||||||
Db   180 agagatggagtttaactccatgttggtcaggctgg................ttttttagt 239

Qy   240 ccctccgtctccggctcccaaagtgct--agatacagagactggccacacatgccgg-ct 295
         ||||||||||||||||||||||||||||||||||||||||
Db   240 ctccctgctcccggctcgctgccccagagaactggccacacatgccggccct 299

Qy   296 ctgcctggctcaattcttgtggtagaacagggttcactgatg-tgcccaagctggtctc 354
         |||||||||||||||||||||||||||||||||||||||||
Db   300 ctgcctggctaattttgtggtagaaacaggcgttcactgatg-tgcccaagctggtctc 359

Qy   355 ctgagctcaagacagtcacacctgcctcagcctcccaaagtgctgggattacaggcgtgcag 414
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   360 ctgagctcaagacagtcacacctgcctcagcctcccaaagtgctgggattacaggcgt-cag 418

Qy   415 ccgtgctggcctttattttttagagacagagtgtcccactcttaccagg 474
         ||||||||||||||||||||||||||||||||||||||||||||
Db   419 ccgtgcctggcctttattttttaaagacacaggtgtgccactcttaccagg 478

Qy   475 atgaagtgcagtggtgtgatcacagctcactcactcctgagatcaagc-at 533
         |||||||||||||||||||||||||||||||||||||||||||||
Db   479 atgaagtgcagtggtgtgatcacagctcactcactgctccagcctccatc 538

Qy   534 ccctcctgcctcagcctcccaaagtagctgggactaggacatgcgaccaccacctgcta 593
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   539 cctcccgcctcagcctccaaagtgctggggacataggcaccaccacacacctgg-TA 597

Qy   594 atttttattttattttgagagacagagtctcaactctgtcacccaggctgga 653
         |||||||||||||||||||||||||||||||||||||||||||
Db   598 atttttattttattttttgagacagagtctc--actcgtcacccaggctgga 656

Qy   654 gtgcagtggcgcaatctcttggctcactgcaacctctgcctcccggttcaagtgattattcc 713
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   657 gtgcagtggcgcaatctcggctcactgcaacctccgcctcccgggttcaagtgattattcc 716

Qy   714 tgcccaggcctcctgagtagctgggactacaggcgcccaccacagctaattttt 773
         ||||||||||||||||||||||||||||||||||||||||
Db   717 tgcccaggcctcctgagtagctgggattacaggcgcccaccacagcctaattttt 776

Qy   774 gtatttttagtagagatgggg--ttcacaccatgttccagttgatctggacc 832
         ||||||||||||||||||||||||||||||||||
Db   777 gtatttttagtagagatggggtttcaccatgttggccagttgatctggacc 836

Qy   833 ttgtgatctgcctgcctcggcct--cccaaagtgctgggattacaggcgtgagccaccacg 891
         ||||||||||||||||||||||||||||||||||||||||
Db   837 ttgtgatctgcctgcctcggcctcccaaagtgctgggattacaggcgtgagctaccacacg 894

Qy   892 cccggcctatttaattttgtttgatttgaaatggaatcactacctgttaccaggctg 951
         |||||||||||||||||||||||||||||||||||||||||||
Db   895 gccggcctatttaattttgtttgttttgaaatggaaacctactcactgttaccaggcg 951

Qy   952 gagtgcaatggccaaatcctcggtcacctgcaacctcccgggctcaagcgattct 1011
         ||||||||||||||||||||||||||||||||||||||
Db   952 gagtgcaatggccaaatcctcggtcacctgcaacctcccgggctcaagcgattct 1011
```

```
Db   955 gagtgcaatgg-caaatctcggtcactcgggctacticgcaaccctctgcctcccggg-tcaagcgattct 1012

Qy  1012 cctgtctcagcctcccaggcagctgggattacgggcacctgccacacacccgcctaatt 1071
         |||||||||||||||||||||||||||||||||||||||||||||||
Db  1013 cctgtctcagcctcccaggcagctgggattacgggcacctgccacacacccggctaatt 1070

Qy  1072 tttgtattttcattagaggggggttcaccatatttgtcaggct-ggtctcaaactcct 1130
         |||||||||||||||||||||||||||||||||||||||||||||
Db  1071 tttgtattttcattagagtgggggttcaccatatttgtcaggct-ggtctcaaactcct 1128

Qy  1131 gacctcaggtgacccacctgccctcagcctcccaaagtgctgggattacaggcgtgagcca 1190
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1129 gacctcaggtgacccacctgcctcagcctcccaaagtgctggggattacaggcgtgagcca 1188

Qy  1191 cctcacccagccgcctaatttttagatataaaaaatatgtagcaatgggggtcttgctgtt 1250
         |||||||||||||||||||||||||||||||||||||||||||
Db  1189 cctcacccagccgcctaatttggaataaaaaataaatatgtagcaatgggggg--tctgctatgt 1246

Qy  1251 tgcccaggctggtctcaaactcctggcttcatgcaatcctcccaaatgggggg--tctgctatgt 1310
         |||||||||||||||||||||||||||||||||||||||||
Db  1247 tgcccaggctggtctcaaactcctggcttcatgcaatcctcccaaatgggggg--tctgctatgt 1306

Qy  1311 cagccagtgcacatttttaaacagttacatcttttatttagtatactagaaagtaataca 1370
         ||||||||||||||||||||||||||||||||||||||||
Db  1307 cagccagtgcacatttttaaacagttacatcttttatttagtatactagaaagtaataca 1366

Qy  1371 ataaacatgtcaaac 1385
         |||||||||||||||
Db  1367 ataaacatgtcaaac 1381
```

```
Query Match   74.98%;   Score 1080.2;   DB 5;   Length 1381;
Best Local Similarity   94.48%;   Pred. No. 4.8e-292;
Matches 1317;   Conservative   0;   Mismatches   53;   Indels   25;   Gaps   18;

Qy      2  tttttttttttgagatgagttgagtttcgtctcttgttgcccaagctggagtgagtgcaatggcgcaat
           ||||||||||| |||||||| || ||||||||||||||||| ||||| ||||| |||||||||||||||||
Db      1  TTTTTTTTTTTGAGATGAGAGTTGTCTCGTCTCTTGTTCGGCTCAAGCGTGGGAGTGCAATGGCGCAAT    60

Qy     62  ctcagctcaccgcaacctccgccctccccgggttcaagcgattctcctgcctcagcctcccc          61
           |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||| ||||
Db     61  CTCAGCTCACCGCAACCTCCGCCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCCC         121

Qy    122  agtagctgggattacaggcatgtgccaccacgctcggtgctaatttttgtatttttttagt         121
           |||||||||||||||| |||| ||||||||||| ||||||||||||| ||||||||||||||
Db    121  AGTAGCTGGGATTACAGGCATGCA-CCACGCTCGGGTAATTTGTATTTTTTTTTTTTTAGT          179

Qy    182  aggagatggaagttt--ctccatgttggtcaggctggtctctgaactcccgacctcaggat         239
           ||||||||||||||    ||||||||||||| ||||||||||||||||||||||||||||||
Db    180  AGAGATGGAAGTTTAAACCTCCATGTTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGATGAT      239

Qy    240  ccctccgtctcggcctccaaaagtgct---agatacaaggactggccacacatgccgg-ct         295
           |||||||||||||| |||||||||||     ||||||||||||| ||||||||||||| ||
Db    240  CTCCCGTCTCGGCCTCCCAAAGTGCTGAGATTACAGGCCACCATGCCGGGCTGGTCTC         299

Qy    296  ctggctggctaattttgtggtagaaacagggttcactgatgtgttgcccaagctggtctc         354
           |||||||||||||||||||||| ||||||||| |||||||||| |||||||||||||||||
Db    300  CTGGCCTGGCTAATTTTTGTGTAGAAACAGGGTTTCACTGATGTGTTGCCCAAGCTGGTCTC     359

Qy    355  ctgagctcaagcaagtccacctgcctcagcctcccaaagtgctgggattacaggcgtgcag         414
           |||||||||||||||||||||||||||||||||| ||||||||||||||||||||| |||
Db    360  CTGAGCTCAAGCAAGTCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGT-CAG      418

Qy    415  ccgtgcctggcctttttattttattttataagacacaggtgtaccactcttaccccagg          474
           ||||||||||||| |||||||||||| ||| ||||||||||||||||||||||||||||
Db    419  CCGTGCCTGGCCTTTTTATTTTATTTTTTAAGACAGGGTGTACCACTCTTACCCCAGG        478

Qy    475  atgaagtgcagtggtgtgaatcacagctcagcccttcaactcctggagatcaagc-at         533
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    479  ATGAAGTGCAGTGGTGTGATCACAACTCACTGCAGCCTTCAACTCCTGAGATCAAGCAAT     538

Qy    534  cctcctgcctcagcctcccaagtagctgggaccaaaagacacgcacaccactgaacctgga         593
           |||||||||||||||||||||||||||||||||||| |||||||||||||||| ||
Db    539  CCTCCTGCCTCAGCCTCCCAAGTAGCTGGGACTACAGGCAAAGACATGCACCACTACACCTGG-TA      597

Qy    594  atttttatttttattttttgagacagagtctcaaactctgtcaccaggctgga         653
           ||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db    598  ATTTTTATTTTTATTTTTTGAGACAGAGTCTC-ACTCTGTCACCCAGGCTGGA       656

Qy    654  gtgcagtggcgaatctggctcactgcaacctctgcctcccgggttcaagtgattcttcc         713
           ||||||||||| |||||||||||||||||||||||||||||||| ||||| |||||||||
Db    657  GTGCAGTGGCGCAATCTCGGCTCACTGCAACCTCCGCCTCCCGGGTTCAAGTTATTCTCC       716

Qy    714  tgcccaagctcctgagtagctgggactacaggcgcccaccaccacgcctagctaatttttt         773
           ||||||||||||| |||||||||||||||||||| |||||||||||| |||||||||||
Db    717  TGCCCCAGCCTCCTGAGTAGCTGGGACTACAGGCGCCCACCACCACGCCTAGCTAATTTTTT      776

Qy    774  gtattttagtagagatggggg--ttcaccatgttgaccaggctgatcctctgacc         832
           |||||||| |||||||||||    ||||||||| |||||||||||||||||| |||
Db    777  GTATTTTTAGTAGAGAT-GGGGTTTCACCATGTTGGCCAGGTTGGTCTCGATCTCCTGACC      836

Qy    833  tttgtgatctgcctcgtgcct--cccaaagtgctggggattacaggcgtgagccacacg         891
           |||||||||||| |||||||    |||||||||||||||||||||||||| || ||||||
Db    837  TTTGTGATCTGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAG--GTCGTGACTCCAC      894

Qy    892  ccccggcttaatttttaattttgtttgtttgaaatggaatctcactctgttaccccaggctg         951
           ||||||||||||||||||  |||  |||||||||||||||||||||||||| || |||
Db    895  GCCCGGCCTAATTTTTTAATTTTTGTTTGTTTGAAATGGAATCTCACTCTGTTGTTAGGGCTG     954

Qy    952  gagtgcaatggccaaatctcggctcactgcaacctctgcctcccgggctcaagcgattct         1011
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    955  GAGTGCAATGG-CAAATCTCGGCTCACTGCAACCTCTGCCTCCCGGG-TCAAGCGATTCT     1012

Qy   1012  cctgtctcagcctcccaaagcagctgggattacgggcaactgccaccacaccccgctaatt         1071
           |||||||||||||||||||||||||||||||||| |  ||||||||||||||||||||||||
Db   1013  CCTGTCTCAGCCTCCCAAGCCAGTCGGGATTACGGG-C-ACCTGCACCACACCCCGCTAAATT      1070

Qy   1072  tttgtatttcaattagagacgggggtttcaccaatgttggtcaggct-ggtctcaaactcct         1130
           |||||||||||||||||||||||||||| || |  |||||||||||   ||||||||||||
Db   1071  TTTGTATTTCATTAGAGACGGGGGTTT--ACCATATTTGTCAGGCTGGTCTCAAAACTCCT      1128

Qy   1131  gacctcaggtgacccaacctgcctcagccctcccaaagtgctgggattacaggcggtgagcca         1188
           ||||||||| ||||||||||||||||||||| |  ||||||||||||||||||||||| |||||
Db   1129  GACCTCAGGTGACCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCA      1190

Qy   1191  cctcaccaccgccggctaattagataaaaaaatatgtagcaggtggggggtcttgctatgt         1250
           |||||||||||| ||||||||||||||||  || |||||||||||||||||||||||||||
Db   1189  CCTCACCACCGCCGGCTAATTTTGAATAAAAAATATGTAGCAATGGGG--TCTGTATGT      1246

Qy   1251  tgcccaggctggtctccaaaacttctgcttcatgcaatcctccaaatgagccacacaacacc         1306
           |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db   1247  TGCCCAGGCTGGTCTCCCAAACTTCTGGCTTCAGTCAATCCTCCAAATGAGCCACACAACACC     1310

Qy   1311  cagccagtcacattttttaaacagttacatcttttatttagtactagagaaagtaataca         1366
           ||||||||||| ||||||||||||||||| ||  |||||||||||||||||||||||||||
Db   1307  CAGCCAGTCACATTTTTTAAACAGTTACATCTTTTATTTTAGTATTAGTACTAGAAAGTAATACA     1370

Qy   1371  ataaacatgtcaaac         1385
           |||||||||||||||
Db   1367  ATAAACATGTCAAAC        1381
```

RESULT 9
US-08-975-080-35
; Sequence 35, Application US/08975080
; Patent No. 6245523
; GENERAL INFORMATION:
; APPLICANT: Altieri, Dario C.
; TITLE OF INVENTION: SURVIVIN, A PROTEIN THAT INHIBITS
; TITLE OF INVENTION: CELLULAR APOPTOSIS, AND ITS MODULATION
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP
; STREET: 1800 M Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036-5869
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/975,080
; FILING DATE: 20-NOV-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/031,435
; FILING DATE: 20-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Adler, Reid G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 044574-5022-01-WO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-467-7000
; TELEFAX: 202-467-7176
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14796 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-975-080-35

Query Match
Best Local Similarity 33.0%; Score 475.4; DB 4; Length 14796;
Matches 877; Conservative 69.8%; Pred. No. 7.9e-123; Indels 59; Gaps 15;
  Conservative 0; Mismatches 321;

```
Qy      1 tttttttttgagatggagtggtttcgtcttgttgcccaggctggagtgcaatggcgcaa 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5380 tttttttttgagatggagtgg-tttcactcttgttgcccaggctggagtgcaatggtgcaa 5438

Qy     61 tctcagctcaccgcaacctccgcctcccgggttcaagcgattctcctgcctcagcctccc 120
          ||| |||||||||||||||||||||||||||||||||||||||||  |||||||||||||||
Db   5439 TCTTGGCTCACTGCAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCC 5498

Qy    121 cagtagctgggattacaggcatgtgccaccacgctcggctaattttttgtatttttttag 180
          |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db   5499 AAGTAACTGGGATTACAGGCATGTGCCACCACGCCCAGGCTAATTTT---TGTATTTTTAG 5555

Qy    181 tagagatgggagtttctccatgttggtcaggctggtctcgaactcccgacctcagatgatc 240
          |||||||||||||||||||||||||||||||||||||||  ||||||||||
Db   5556 TAGAGATGGGGTTTCACCATGTTGGCCAGGCTGGTCTTGAACTCCTGACCTC--GTGATT 5613

Qy    241 cctccgtctcggcctcccaaagtgctagatacaggactggcacacatgcccggcgtctgcc 300
          ||||||||||||||||||||||||||||||||||||||||    |||||||||||||||||
Db   5614 CGCCCACCTTGGCCTCCCAAAGTGCTGG---GATTACAGGCGTGAGCCACCGCGGCTGGCC 5670

Qy    301 tggctaattttgtggtagaaacaggggttcactcatgttgcccaagtctgggtctcctg--- 357
          |||||||||||||||||||||||||||| |||||||||||||||||||||
Db   5671 TTTTTTTTTTTTTTGTTCT-GAGACACAGTTTCACTCTGTTACCCAGGCTGGAGTAGGTGG 5729

Qy    358 --agctcaagcagtccaactgcctccagcctcccaaagtgctgggattacaggcgtgcagc 415
               |||||||||||||||||||||||||||||||||||||||||||||
Db   5730 CCTGATCTCGGATCACTGCAACCTCCGCCTCCCGGCTCCCAAGTGATTGGGATTACAGG 5789

Qy    416 cgtgcctggcctttttattttatttttttaagacaggtgtcccactcttaccagga 475
          |||||||||||  ||||||||||||||||||||||||||||||||||||||
Db   5790 CTCCCAAGTAGCCAGAGTTACAGGCATGCGCCACCACACCAGGTAATTTTGTATTTTT 5849

Qy    476 tgaagtgcagtggtgtgatca--cagctcactgagcctcaactgagatcaag-ca 532
          |||||||||||||||||||||   ||||||||||||||||||||||||| ||
Db   5850 GGTAGAGACGAGGTTTCACCATGTTGGCCAGGCTGGTCTTGAACTCCTGACCTCAGGTGA 5909

Qy    533 tcctcctgcctcagcctcccaagtgctgggaccaaagacatgcaccacctggc- 591
          ||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db   5910 TCCACCCGGCTCACGTCCCAAAGTGCTGAGATTAPAGGTGTGAGCCACCACGTGA 5909

Qy    592 -----------taattttattttttttaattttttgagacagagtctcaactctgt 640
                     || ||||||||||||||||||||||||||||||||
Db   5970 TCAGGAAGTATATTTTTTATTTTTATATTTAAATTTATTTTGAGATGGAGTCT-TGCTCTGT 6028

Qy    641 cacccaaggctggagtgcagtgggcgcaatcttggctcactgcaacctcgcctcccggtt 700
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db   6029 CGCCCAGGCTAGAGTGCAGGCGGATCTCGGCTCACTGCAACCTCCGCCCCCAGGTT 6088

Qy    701 caagtgattctcctgcctcagcctcccgagtagctggactacaggcgc---ccaccacg 757
          |||||||||||||||||||||||||||| |||||||||||||||||   |||||||||
Db   6089 CAAGCCATTCTCCTGCCTCAGCCTCCGAGTAGCTGGGACTACAGGCGCCCGCCACCACA 6148

Qy    758 cctagctaatttttttgtattttttagtagagatggg-gttcaccatgttcgccaggttga 816
          |||||||||||||||||||||||||||||||||||  ||||||||||||||||||
Db   6149 CCCGGCTAATTTTTTGTATTTTTAGTAGAGACGGGGTTTCACCATGTTAGCCAGGAGG 6208

Qy    817 tcttgatctctggacctgtgatctcgctccgcctcccaaagtgctgggattacag 876
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   6209 TCTTGATCTCCTGACCTCGTGATCTGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAG 6268

Qy    877 gcgtgagccacacgcccggcgcttatttttaatttttgttttgttttgttgttagg 937
          |||||||||||||||||||  |||||| |||||||||||||
Db   6269 GTGTGAGCCACCACCACCCGCCT-------ATTTTAITTTTTTTTGTTTGTTTTTGTTGATT 6320

Qy    937 ctgttacccaggctggagtgcaaatgccaaatcctggctcactgcaacctctgcctcccg 996
          ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db   6321 CTGTCACCTGGGCTTCGCAGTGCACTGCAGCCTCGAACTCCTG 6379
```

RESULT 10
US-09-630-706-10
  Sequence 10, Application US/09630706
  Patent No. 6277640
  GENERAL INFORMATION:
  APPLICANT: C. Frank Bennett
  APPLICANT: Lex M. Cowsert
  TITLE OF INVENTION: ANTISENSE MODULATION OF HER-3 EXPRESSION
  FILE REFERENCE: RTS-0053
  CURRENT APPLICATION NUMBER: US/09/630,706
  CURRENT FILING DATE: 2000-08-01
  NUMBER OF SEQ ID NOS: 94
  SEQ ID NO 10
    LENGTH: 14796
    TYPE: DNA
    ORGANISM: Homo sapiens
    FEATURE:
      NAME/KEY: CDS
      LOCATION: (2811)...(2921)
      NAME/KEY: CDS
      LOCATION: (3174)...(3283)
      NAME/KEY: CDS
      LOCATION: (5158)...(5275)
      NAME/KEY: CDS
      LOCATION: (11955)...(12044)
US-09-630-706-10

Query Match
Best Local Similarity 33.0%; Score 475.4; DB 4; Length 14796;
Matches 877; Conservative 69.8%; Pred. No. 7.9e-123; Indels 59; Gaps 15;
  Conservative 0; Mismatches 321;

```
Qy    997 ggctcaagcgattctcctgcctcagcctcccaagcagtgggattacaggcaacctgccac 1056
          |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6380 AGCTCAAGCGAGATCCTCCCACCTCAHTCCCTTCACAAAGTAATTGGGGACTACAGGTGCACCCCAC 6439

Qy   1057 cacccccgctaattttgtat-------------tttcattagaggcggggtttcac 1101
          || |||||||||||||||                |||||||||||||||||||
Db   6440 CATGCCCACCTAATTTATTTATTTATTTATTTTTTCATPAGAGATGAGGTTTCC 6499

Qy   1102 catatttgtcaggctggtctcaaaactccttgacctcaggtgacccacctgcctcagcctc 1161
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6500 TGTGTGTGCCAGGCTGGTCTTGAACTCCTGAGCTCTACGCGGGGATCCTTTTGCCTGGGGCTCC 6559

Qy   1162 caaagtgctgggattacaggcgtgagccaccacgccggctaatttagataaa 1218
          || |||||||||||||||| |||| || ||||| |||||| ||||||
Db   6560 CAAAGTGCTGAGATTACAGGCATGAGCCACCACGGTGCCCAGCTAGGAACTCATTTTTAAA 6616
```

```
Qy      1 tttttttttgagatggagtggtttcgtcttgttgcccaggctggagtgcaatggcgcaa 60
          ||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db   5380 tttttttttgagatggagtgg-tttcactcttgttgcccaggctggagtgcaatggtgcaa 5438

Qy     61 tctcagctcaccgcaacctccgcctcccgggttcaagcgattctcctgcctcagcctccc 120
          ||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5439 tcttggctcactgcaacctccgcctcccgggttcaagcgattctcctgcctcagcctccc 5498

Qy    121 cagtagctgggattacaggcatgtgccaccacgctcggctaattttttgtatttttttag 180
          |||||||||||||||||||| ||||||||||||||||||  ||||||||||||
Db   5499 aagtaactgggattacaggcatgtgccaccacgcccaggctaatttt---tgtatttttag 5555

Qy    181 tagagatgggagtttctccatgttggtcaggctggtctcgaactcccgacctcagatgatc 240
          |||||||||| ||||||||||||||||||||||||| |||||||||||
Db   5556 tagagatgggggtttcaccatgttggccaggctggtcttgaactcctgacctc--gtgatt 5613

Qy    241 cctccgtctcggcctcccaaagtgctagatacaggactggcacacatgcccggcgtctgcc 300
          ||||||||||||||||||||||||||||   ||||||||||  |||||||||||||||
Db   5614 cgcccaccttggcctcccaaagtgctgg---gattacaggcgtgagccaccgcggctggcc 5670

Qy    301 tggctaattttgtggtagaaacaggggttcactcatgttgcccaagtctgggtctcctg--- 357
          |||||||||||||||||||||||||||| |||||||||||||||||||||
Db   5671 ttttttttttttttgttct-gagacacagtttcactctgttacccaggctggagtaggtgg 5729
```

```
RESULT 11
US-09-496-694B-3
Sequence 3, Application US/09496694B
Patent No. 6335194
GENERAL INFORMATION:
  APPLICANT: C. Frank Bennett
  APPLICANT: Elizabeth J. Ackermann
  APPLICANT: Eric E. Swayze
  APPLICANT: Lex M. Cowsert
  TITLE OF INVENTION: ANTISENSE MODULATION OF SURVIVIN EXPRESSION
  FILE REFERENCE: ISPH-0439
  CURRENT APPLICATION NUMBER: US/09/496,694B
  CURRENT FILING DATE: 2000-02-02
  PRIOR APPLICATION NUMBER: 09/286,407
  PRIOR FILING DATE: 1999-04-05
  PRIOR APPLICATION NUMBER: 09/163,162
  PRIOR FILING DATE: 1998-09-29
  NUMBER OF SEQ ID NOS: 249
SEQ ID NO 3
  LENGTH: 14796
  TYPE: DNA
  ORGANISM: Homo sapiens
  FEATURE:
  NAME/KEY: CDS
  LOCATION: (2811)...(2921)
  NAME/KEY: CDS
  LOCATION: (3174)...(3283)
  NAME/KEY: CDS
  LOCATION: (5158)...(5275)
  NAME/KEY: CDS
  LOCATION: (11955)...(12044)
US-09-496-694B-3

Query Match      33.0%;  Score 475.4;  DB 4;  Length 14796;
Best Local Similarity   69.8%;  Pred. No. 7.9e-123;
Matches  877;  Conservative   0;  Mismatches  321;  Indels  59;  Gaps  15;
```

(alignment data — Qy/Db sequence rows from position 1/5380 through 6029/6088, shown in two columns of the original page)

```
RESULT 12
US-09-813-817-3/c
; Sequence 3, Application US/09813817
; Patent No. 6340583
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001178
; CURRENT APPLICATION NUMBER: US/09/813,817
; CURRENT FILING DATE: 2001-03-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 59065
;   TYPE: DNA
;   ORGANISM: Human
US-09-813-817-3

Query Match             28.9%;
Best Local Similarity   67.7%;
Matches  815;  Conservative    0;  Mismatches  326;  Indels   63;  Gaps   14;
Score 416.4;  DB 4;  Length 59065;
Pred. No. 4.3e-106;
```

```
Qy    701 caagttattctctgctgcccagcctcctgagttagtctgggactacaggcgc---ccaccacg 757
          |||||| ||||| ||||| || ||||||| ||  ||||||||||||||||||||    ||||| ||
Db   6089 caagccattcctctgcctgccccagcctcccagctagtctgggactacaggcggccgccacaca 6148

Qy    758 cctagctaatttttgtatttagtagagatggg-gttcaccatgttgccaggttga 816
          || ||||||||||| ||||| ||||||| ||| |||||||||||| ||||| | ||
Db   6149 cccggctaattttttgtatttttagtagagacgggtttcaccatgttagccaggagg 6208

Qy    817 tcttgatctctggaacttgtgatctgcctcggcctcccaaagtgctggggattacag 876
          |||||||||| ||| || |||||||||| ||||||||||||||||||| |||||||||
Db   6209 tcttgatctcctgacctcgtgatctgcccgcctcggcctcccaaagtgctgggattacag 6268

Qy    877 gcgtgagccacagcgccggcctattttttaatttttgtttgttgaaatgtcact 936
          | |||||||||  |||||||| |||||| ||||||| |||  ||||| | |||||
Db   6269 gtgtgagccacaccacccggct------attttattttttgagacagggactcact 6320

Qy    937 ctgttacccaggctggagtgcaaatctcggctcactgcaactctcgcctccog 996
          ||||| || |||||| |||||| | |  ||||||||| |  |||| ||||||| 
Db   6321 ctgtcacctgggctgcagtgcagtgg-tacacatagctcactgcagccttgaacctcg 6379

Qy    997 ggctcaagcgatctcctgtcctcagccctccaagcagctgggattacaggcgcacctgccac 1056
          ||||||| ||| |||| | |||| || ||| ||||   |||| ||||||||| || | ||||
Db   6380 agctcaagtgatccctcccacctcatcctcaccaagtaattgggactacaggtgcaccccac 6439

Qy   1057 cacacccccgctaatttttgtat----------tttcattagagacggggtttcac 1101
          ||  ||| | || ||||| |||           ||  ||||||| ||| |||||| 
Db   6440 catgcccaactaattatttattattttattttttattagaggatgaggtttcc 6499

Qy   1102 catattgtcaggctggtctccaaactccgacctcaggtgatccacctgcctcagcctc 1161
          | ||||| ||||||||| |  |||| ||||| |||| ||||||  ||| || ||||| 
Db   6500 tgtgttgtccaggctggtcttgaactcctgacactcacgtgatcctttgcctgggcctc 6559

Qy   1162 caaagtgctgggattacaggcgtgagccacctcacccacgcccggctaattttagataaa 1218
          |||||||||| |||||||||| |||||||||  | ||| ||||||||| ||  ||||||
Db   6560 caaagtgctgagattacaggcatgagccaccgtgcccggctaggaatcattttaaa 6616
```

```
Qy     42 ctggagtgcaatggcgcaatctcaccgcacctccgcctccgggttcaagcgat 101
          |||||| | ||||||||  | ||| |||| |||| |||||||||||||||| ||
Db  15163 CTGGAGGGGCAGTGGCGCGGGATCCTCAGCACACTCCCATCTCCAGGTTCAAGCTAT 15104

Qy    102 tctcctgcctcagcctcccagtagctgggattacaggcgatgtgcaccacggctccggcta 161
          |||||||||| ||||||| |||||||||||||||||| | |||| ||| | ||| |||| 
Db  15103 TCTCCTGCCTTAGTCTCCTAAGTAGCTGGGATTACAGGTGCCTGCCACCATCGTGGCTGGCTA 15044

Qy    162 attttgtattttttttagtagagatggaggggttctccacgttgggctggtctccgaa 221
          ||||||||||||||||| |||||| ||||| |||||||||| |||| || || ||| 
Db  15043 atttttgtattttttttagtagagatggaggtgttctccacgttggctggtctccga 13971
```

```
Db  15043 ATTTT----TGTAAATTTTAGTAGAGACAAGGTTTTCACCATGTTGGCCAGGCTGGTCTGGAA 14987
Qy    222 ctccgacctcagatgatcctcccacctcggcctcccaaagtgct-agatacagaggactgg 280
          |||| |||||| |||||| |  |||| |||||||||||||||||  || || ||||| || 
Db  14986 CTCCTGAACCTCAGGTGATCTCCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGCG 14927

Qy    281 caaccaatgccggtcctcgcctggctaatttttgtgtttagtaaaacagggtttcactgatgtg 340
          || ||| ||||| |   | |||||||   ||||||| |||| ||  ||| |||||| |||||
Db  14926 AGCCACCGCGCCCAACCTCCATTTTTTTTTTTTTTTGGTAGACAGAGTCTTCTCTGTCG 14867

Qy    341 cccaagctg---------------------gtctcctgagctca 363
          ||||| |||                     || ||||||| || 
Db  14866 CCCAGGCTGGAGTGCAGTGGCGCGGTCTCAGCGTCACTGCAAGCTCCGCCTCCCGGGTTCA 14807

Qy    364 agcagtccacctgcctcagcctcccaaagtgctgg-gcgtgcagcgtgcc 421
          | | |  |||||||||||||||| |||||||||||  |  |||| |||||
Db  14806 CGCCATTCTCCTGCCTCAGCCTCTGAGTAGCTGGGACTACAGGCGCCCGCCACCATGCC 14747

Qy    422 tggcctttttatttttattttttttaagacacagcagtgtcccactcttaaccaggatgaagt 481
          |||||| ||| ||||||||| |  |||||||| ||| |||  ||| | | ||||||||| | |
Db  14746 TGGCCTAATTTTGTGTATTTTTAGTAGAGACGGGGTTTCACTGTGTTAGCCAGGATGGTCT 14687

Qy    482 gcagtggtgatcaacgtcactgcgagcctcaactcctgagatcaaagcaatcctcctgc 541
          ||| |  | ||| | | | || ||| ||||| |||||  || ||||| ||||| ||| 
Db  14686 -CGATCTCCTGACCTTGTGATCCGCCCGCCTCCCAAAGTGCTGGGATTACAGGC 14628

Qy    542 ctcagcctcccaagtagctgggaccaaaga-----catgcaccactacacctggctaatt 596
          |||||||||| ||| |||||||| | ||||     ||||| ||||| ||| ||||||| |
Db  14627 GTGAGCCACCGCGCCCTGCCTGGCCTCCATATTTTTTTTTATTTTATTTTATTTTATTTT 14568

Qy    597 tttattttttattttttttttgagacagagtctcaactctgtcacccaggctgggagtg 656
          |||||||| ||||||||  | ||||||||||||| ||||||||| |||||||||| |||
Db  14567 TATTTTATGTTTATTTATTTTTATTGAGACAGAGTCTC-GCTCTGTCTCCCAGGCTGGAGTG 14509

Qy    657 cagtggcgcgaatcttggctcactgcaacctctgcctcccgggttcaagttattcctgc 716
          ||||| |  | || ||||| |||||||||| | ||||| |||||||||| ||| ||||
Db  14508 CAGTGACATGCATGTTGGCCCACTGCAACCTCCGCCTTCCCGGGTTCAAGCGATTCTCGTGC 14449

Qy    717 cccagcctcc---tgagtagctgggactacaggcgc---ccaacaagcctagctaatttt 770
          ||||||||||   | ||||||||||||||||||||||    ||| ||| |||||||| |
Db  14448 CTCAGCCTCCCAAGTGCTGAGTAGCTGGGACTACAGGCGCCCACCAACCAGCTAATTT 14391

Qy    771 tttgtattttagtagagatggg-ttcaccatgttcgccaggttggtctgatctctgg 829
          |||||||| |||||||| ||||  ||||||||||| |||||| |||||| | |||| 
Db  14390 TTTGTATTTTCAGTTGAGATGGGGTTTCACCATGTTGCCAGGCTGGTCTCGAATCTCCTG 14331

Qy    830 acct--tgtgatctgcctcggcctcccaaagttgctgggattacaggcgtgagccac 887
          ||||  ||||||| ||||||||||||||||| ||||||||||||| |||||||| ||
Db  14330 ACCTCAAGTGATCACCTGCCTCGGCCTCCCAAAGTGCTGGGATTACTGGCGTGAGCAA 14271

Qy    888 cacgcccggctattt-ttaatttttgttttgtggaatctcaactctgttaccca 946
          ||||||  || | ||  |  |||||||||||| || | || ||||||| |||||
Db  14270 CGCGCCTGGGCTCGTTTCTTTTTTTTTTTTTTTTGAGATAGGGTCTTGCCTCTGTCATCCA 14211

Qy    947 ggctggagtggcaatggccaaaatctggtcactgcaacctctgcctccgggttcaaagc 1006
          ||||||  |||| ||||    ||| || ||||||||| ||| ||||| |||||||||| 
Db  14210 GGCTGGCATGTAGTGGTGCGATCACGGCTCACTGCAGCCTCAACCTCCTGGGCTCAAGCA 14151

Qy   1007 attctcctgtctcagcctcccaagcagctgggattacgggcaacctgccacacacccc 1066
          |||||||| ||||||||||| ||| |||||||||||| |||  ||| |||||||||| 
Db  14150 ATCCTCCTGCCCCAGCCTCCCGAGTAGCTGGGATTACAGGCGCCCAGCACACACACCCGGC 14091

Qy   1067 taattttgtattttcattagaggcgggggtttcaccatatttgtcaggctggtc-tcaaa 1125
          || ||||||||||| ||||||| |||  | |||||||| ||||||||||||||  |||||
Db  14090 TAAATTTTTGATTTTTTGTGGGGGATAGGAGGTCTTGCTGTCTTACCCAGACTGAICTGAAA 14031

Qy   1126 ctcctgacctcaggtgacccaactgcctcagccttccaaagtgctgggtgggattacaggcgtg 1185
          |||||| ||||| |   |  ||| ||||||||| | ||||||||||  || ||||| ||||||
Db  14030 CTCCTGAGTTCAAGCAATCCTCCCGCCTCAGCCTCCCAAAGTGCTGCTGGATTATATTAAGTGTA 13971

Qy   1186 agcc 1189
          ||||
Db  13970 AGCC 13967
```

RESULT 13
US-08-687-080-101/c
Sequence 101, Application US/08687080
Patent No. 5965427
; GENERAL INFORMATION:
; APPLICANT: Gregory Dolganov
; TITLE OF INVENTION: Human RAD50 Gene and Methods of Use Thereof
; NUMBER OF SEQUENCES: 175
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US 08/687,080
; FILING DATE: 17-JUL-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/592,126
; FILING DATE: 26-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 4600-0111.30
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 101:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5543 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ANTI-SENSE: NO
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: 3' END OF INTRON 19 OF RAD50 GENOMIC
; INDIVIDUAL ISOLATE: SEQUENCE
US-08-687-080-101

Query Match   27.8%; Score 400.8; DB 2; Length 5543;
Best Local Similarity 72.1%; Pred. No. 3.4e-102;
Matches 610; Conservative   0; Mismatches 217; Indels 19; Gaps 6;

Qy   592 taattttattttattttaattttttgagacagagtctcaacttcctgtcacccaggctg 651
         | ||||| |||||| ||||| | ||||| ||||||
Db 1274 TGACTTTTTTTTTTTTTTTTTTTTTTTTAAGACGGAGTCT-GGCTCTGTCACCCAGGCTG 1216

Qy   652 gagtgcaagtggcgcaatccttggctcactgcacctccccgggttcaagtgattattct 711
         || ||||||||| || |||||| || ||||||||||||| || |||
Db 1215 GAGTGCAGTGGTGCGGATCTCGGCTCACTGCAAGCTCTGCCTCCCAGGTTCACGCCATTCT 1156

Qy   712 cctgccccagccctcctgagtgagctgggactacaggcgcgccacac---gcctagctaatt 768
         |||| || |||| |||| |||| | ||| |||| |||| |||| |||
Db 1155 CCTGCCTCAGCCTCCTGAGTAGCTGGGACTACAGGCGCCCACCACCACGGCCTGGCTAATT 1096

Qy   769 ttttgtatttttagtagagatgggg--ttcaccatgttgatcttgatctgt 827
         ||| | |||| | |||||||||||||||||||||| || ||| | ||
Db 1095 TTTTGTATTTTTAGTAGAGACGGGGTTTCGGCGTGTTGGCCAGGATGGTCTCGATCTCC 1036

Qy   828 ggaccttgtgatctgcctcggccctcccaaagtgctgggattacagcgtgagccac 887
         ||||| ||||||||||||| |||||||||||||||||||||||||
Db ...

Db 1035 TGACCTCGTGATCTGCGCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCAC 976

Qy   888 cacgcccggcttat---------ttttaatttttgtttgttttgaaatggaatccaactc 937
         |||||| |||| || || || ||| | ||
Db   975 CACGCCCTGGCTGTCTGTTTTTGTGTTTTTTCTGTTTGTTTTGTTTGAGACGGGAGTTTCACTC 916

Qy   938 -tgttacccaggctgagtgcaatgcgcaaatctcggctcactgcaacctctgcctccccg 996
          |||| |||||||| |||| || || ||| || ||||||||| || | ||| ||
Db   915 TTGTCACCCAGGCTGAAGTGCAATGGTGATCTCGGCTCACTGCAACCTCTGCCTCCCA 856

Qy   997 ggctcaagcggattctctccgtctcagcctcccaagcagctgggattacgggcaacctgccac 1056
         || |||||| ||||||| | |||||||| ||||||||||||||||| | |||||||||||
Db   855 GGTTCAAGCGGATTCTCCTCGCCTCAGCCTCCTGAGTACCTGGGATTACAGGCGCGTGTCAC 796

Qy 1057 cacacccccgctaattttttgttatttttcattaggaggcggggttcaccatattgtcaggct 1116
         || ||||| | |||| ||| | ||| ||||||| |||||| |||||||| ||||||||
Db   795 CACACCTGGCTAATTTTTTGTATTTTCAGTAGAGANGGGGTTTTACCATATTGGCCAGGCT 736

Qy 1117 ggtctcaaactcctgaacctcaggtgaccaacctgcctcagccttccaaagtgctgggatt 1176
         ||||||||| ||| | || | |||||||||||| || ||||||||||||||||||| |||
Db   735 AGTCTTGAAACTCCTGACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGAATT 676

Qy 1177 acaggcgtgagccacctcacccagccgcggctaatttagataaaaaatatgtgcaaatggg 1236
         ||||| | | ||| ||| |||| | | |||||||| |||| |||||| | | ||| ||
Db   675 ACAGGCANGAGTCACCTCGCCCGGCCGCCTCGCTAATTTTTTGTATTTTAGTAGAGAATATTG 616

Qy 1237 gggtcttgctatgtgcccaggctggtcttaaacttctggcttcatgcaatcctccaaa 1296
         || ||| |||| || ||||||||||| | | |||||| ||| || || ||||
Db   615 AGCATATCAGACTTCATGCCATACTTTTTATACACTTCT---CTCCTGGTTTCATTTACTTT 559

Qy 1297 tgagccacacaccagccagtcaacattttaaacctgcaaattcagtagtaacagagttctttta 1356
         ||| ||| || |||||||||| | ||| | ||| ||||||||| ||||| ||||
Db   558 CTTGCCCTTATTTCTACACGTCGCCTTGTTTCCCATTAATTTGAACAAGGAAGGTACAATTTATCTTTGC 499

Qy 1357 tagaaagtaatacaataaacatgtcaaacctgcaaattcagtaatactggtcttta 1416
         ||| ||||| | | ||||||| ||| | || || || ||| || |
Db   498 TCTATTGTATATATAACTAAGTAAATAATTTCTGAACAAGGAAGGTRCAAAGTAAACTAA 439

Qy 1417 taactt 1422
         ||| ||
Db   438 TACCAT 433

RESULT 14
US-08-257-963B-9
Sequence 9, Application US/0825796 3B
Patent No. 5840686
; GENERAL INFORMATION:
; APPLICANT: Chader, Gerald J.; Becerra, S.
; APPLICANT: Patricia; Schwartz, Joan P.;
; APPLICANT: Taniwaki, Takayuki
; TITLE OF INVENTION: PIGMENT EPITHELIUM
; TITLE OF INVENTION: DERIVED FACTOR: CHARACTERIZATION OF ITS NOVEL
; TITLE OF INVENTION: BIOLOGICAL ACTIVITY AND SEQUENCES ENCODING
; TITLE OF INVENTION: AND EXPRESSING THE PROTEIN
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan
; STREET: 345 Park Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/257,963B
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:

Left column

APPLICATION NUMBER: 07/952,796
FILING DATE: 24-SEPT-1992
ATTORNEY/AGENT INFORMATION:
NAME: DOROTHY R. AUTH
REGISTRATION NUMBER: 36434
REFERENCE/DOCKET NUMBER: 20264126US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 4421 Base Pairs
TYPE: Nucleic Acid
STRANDEDNESS: Double
TOPOLOGY: Unknown
MOLECULE TYPE: Genomic DNA
ORIGINAL SOURCE:
ORGANISM: Human
FEATURE:
NAME/KEY: JT101
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION: 7.1 kb Bam HI
OTHER INFORMATION: fragment Derived from human placental
OTHER INFORMATION: genomic DNA
US-08-257-963B-9

Query Match
Best Local Similarity 27.3%;   Score 393.6; DB 2;   Length 4421;
Matches 512; Conservative 0; Mismatches 94; Indels 12; Gaps 5;
82.8%;   Pred. No. 3.1e-100;

```
Qy   596 tttattttattttaatttttgagacagagtctcaactctgtcacccaggctggagt   655
Db  3485 TCTTAGTATTTTTTTTTTTTTCTTTGAGATGGAGTCT-GGCTCTGTGCACCCAGGCTGGAGT 3543

Qy   656 gcagtggcgcaatcttggctcactgcaacctctgcctcccgggttcaagtgattattcctg 715
Db  3544 GCAGTGGCGCGGGATCTCGGCTCACTGCAAGCTCTTCCTCCCGGGTTCACCCAATTCTCCTG 3603

Qy   716 ccccagcctcctgagtagctgggactacaggcg--ccacccacgcctagctaattttt    772
Db  3604 CCTCAACCTCCCGAGTAGCTGTGAATACAGGCGTCCGCCACCACGCCCGACTAATTTTTTT   3663

Qy   773 tgtattttagtagagatggg-ttcaccaatgttccgccaggttgatcttctgac       831
Db  3664 TGTATTTTTAGTAGAGACGGGGTTTCACCGTGTTAGCCCAGGATGGTCTCGATCTCCTGAC  3723

Qy   832 cttgtgatctgcctgcctcccaaagtcgcttggattacaggcgtgagccaccacg        891
Db  3724 CTCGTGAGATCTGCCTGCCTCGGCCTCCCAAAAGTGCTGGGATTACAGGCGTTGAGCCACCGCA 3783

Qy   892 ccgggcttattttgtttgtttgaaatggaatctcaactctgttacccaggctg         951
Db  3784 CCCGGCCTCTTATTTTTTTTT-----TTTTGAGATGAGGAGTCTCACACTGTCACCTGGGCCTG 3837

Qy   952 gagtgcaaatgcccaaaatctcggctcactgcaacctctgcctcccgggctcaagcgcattct 1011
Db  3838 GAGTGCAGTGGAACGCGANTCTCGGCTCACTGCAACCTCCGCCTTCCTGGGTTCAAGCGAGATTCT 3897

Qy  1012 cctgtctcagcctcccaagcagctgggattacgggcacctgccaccaccacccgccta-at  1070
Db  3898 CCTGCCTCACCCTCCCAAGTAGCTGGGATTACAGGTGCCCACCACCACGCCTGGCTAGTT    3957

Qy  1071 tttttgtattctcattagaggcggggtttcaccatcattgtcaggctggtctcaaactcct  1130
Db  3958 TTTTGTATTTTTTAGTAGAAGATGGGGTTTCACCATGTTGGCCAGGCTGGTCTTGAAACTCCT  4017

Qy  1131 gacctcaggtgacccacctgccagccttccaaagtgctgggattacaggcgtgagcca     1190
Db  4018 GACATCAGGTGATCCGCCCGCCTCAGCCTTCCAAAGTGCTGGGATTACAGGCGTGAGCCA   4077

Qy  1191 cctccaccagccggctaa    1208
```

Right column

```
Db  4078 CCATACCTGGCCAGCAAA 4095
```

RESULT 15
US-08-367-841A-9
Sequence 9, Application US/0836784A
Patent No. 6319687

GENERAL INFORMATION:
APPLICANT: Chader, Gerald J.; Rodriguez,
APPLICANT: Ignacio R.; Mazuruk, Krzysztof;
APPLICANT: Tombran-Tink, Joyce
TITLE OF INVENTION: PIGMENT EPITHELIUM
TITLE OF INVENTION: DERIVED FACTOR: CHARACTERIZATION GENOMIC
TITLE OF INVENTION: ORGANIZATION AND SEQUENCE OF THE PEDF GENE
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan & Finnegan
STREET: 345 Park Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/367,841A
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
FILING DATE: 30-DEC-1994
APPLICATION NUMBER: 08/257,963
FILING DATE: 07-JUN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/952,796
FILING DATE: 24-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: DOROTHY R. AUTH
REGISTRATION NUMBER: 36434
REFERENCE/DOCKET NUMBER: 20264126US2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 4421 Base Pairs
TYPE: Nucleic Acid
STRANDEDNESS: Double
TOPOLOGY: Unknown
MOLECULE TYPE: Genomic DNA
ORIGINAL SOURCE:
ORGANISM: Human
FEATURE:
NAME/KEY: JT1
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION: 7.1 kb Bam HI
OTHER INFORMATION: fragment Derived from human placental
OTHER INFORMATION: genomic DNA; Also referred to as JT101
US-08-367-841A-9

Query Match
Best Local Similarity 27.3%;   Score 393.6; DB 4;   Length 4421;
Matches 512; Conservative 0; Mismatches 94; Indels 12; Gaps 5;
82.8%;   Pred. No. 3.1e-100;

```
Qy   596 tttattttattttaatttttgagacagagtctcaactctgtcacccaggctggagt   655
Db  3485 TCTTAGTATTTTTTTTTTTTTCTTTGAGATGGAGTCT-GGCTCTGTGCACCCAGGCTGGAGT 3543

Qy   656 gcagtggcgcaatcttggctcactgcaacctctgcctcccgggttcaagtgattattcctg 715
```

```
Db  3544 GCAGTGGGCGGCGGATCTCGGGCTCACTGCAAGCTCTTCCTCCCGGGTTCACCCCATTCTCCTG 3603
              |||||||||||  ||||  ||||||||  |||  |||||  |||   ||||||||||

Qy   716 cccagcctcctgagtagctgggactacaggcg---cccaccaagcctagctaatttt 772
              |||  ||||  ||||| ||||||||||||||||      |||||||||||||||||||
Db  3604 CCTCAACCTCCCGAGTAGCTGTGAATACAGGCGTCCGCCACCACGCCCGACTAATTTTTTT 3663

Qy   773 tgtatttttagtagagatgggg-ttcaccatgttcgccaggttgatcttctggac 831
              |||||||| |||||||  |||| ||||| ||||| |||||| || |||  |||||
Db  3664 TGTATTTTTTAGTAGAGACGGGGTTTCACCGTGTTAGCCAGGATGGTCTGGAATCTCCTGAC 3723

Qy   832 cttgtgatctgcctcggcctcccaaagtgctgggattacaggcgtgagccaccacg 891
              ||| |||||||||| |||  |||||||||||||||||||||||||| |||||| |
Db  3724 CTCGTGATCTGCCTGCCTCGGCGCCTCCCAAAGTGCTGGGATTACAGGCGTTGAGCCACCGCA 3783

Qy   892 cccggcttattttaattttgtttgaaatggaatctcactctgttacccaggctg 951
              |||||  |  |  | || ||||  ||  |  |||||| |||||| ||| ||  |
Db  3784 CCCGGCCTCTTPATTTTTTTT-----TTTTGAGATGGAGTCTCACACTGTCACCTGGCCTG 3837

Qy   952 gagtgcaatggccaaatctcggctcactgcaacctctgcctcccgggctcaagcgattct 1011
              ||||||| ||   |||||||||||||||||||||||  |||||  |||||||  |||||
Db  3838 GAGTGCAGTGGGAGGCGAATCTCGGCTCACTGCAACCTCCGCCTCCTGGGTTCAAGAGATTCT 3897

Qy  1012 cctgtctcagcctcccaagcagctgggattacgggcacctgccacacacccgcta-at 1070
              ||||  |||||||||||||  ||||||||||| |||  |  |||||||| |||  |  ||
Db  3898 CCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCCCACCACCACGCCCTGGCTPAGTT 3957

Qy  1071 ttttgtatttcattagagcgggggtttcaccacatatttgtcaggctggtctcaaactcct 1130
              |||||||||| ||||||| |||||||||||||| |||| |||||||||||||| ||||||
Db  3958 TTTTGTATTTTTAGTAGAGATGGGGTTTCACCATGTTGGCCAGGCTGGTCTTGAACTCCT 4017

Qy  1131 gacctcaggtgacccacctgcctcagcccttccaaagtgctgggattacaggcgtgagcca 1190
              ||| ||||||||| |||| || ||  |||  |||||||||||||||||||||||||| ||||
Db  4018 GACACTCAGGTGATCCGCCCGCCACCTTAGCCTCCCAACTTGCTGGGATTACAGGCGTGACCA 4077

Qy  1191 cctcaccacgccggctaa 1208
              || |||||||||||||||
Db  4078 CCATACCTGGCCAGCAAA 4095
```

OM nucleic - nucleic search, using sw model

Copyright (c) 1993 - 2000 Compugen Ltd.

GenCore version 4.5

Run on: April 24, 2002, 17:05:21 ; Search time 1728.26 Seconds
(without alignments)
17460.374 Million cell updates/sec

Title: US-09-964-678A-1

Perfect score: 1442

Sequence: 1 ttttttttttgagatggag..........ttaacaaagctttagagca 1442

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 1046326293 residues

Total number of hits satisfying chosen parameters: 3595312

Post-processing: Listing first 45 summaries

Minimum Match 0%
Maximum Match 100%

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Database : GenEmbl:*

```
1:  gb_ba:*
2:  gb_htg:*
3:  gb_in:*
4:  gb_om:*
5:  gb_ov:*
6:  gb_pat:*
7:  gb_ph:*
8:  gb_pl:*
9:  gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*
```

SUMMARIES

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Match Length	DB	ID	Description
1	1442	100.0	1442	6	AR051550	Sequence
2	1442	100.0	1442	6	AR072690	Sequence
3	1442	100.0	1442	6	AR073235	Sequence
4	1442	100.0	1442	9	AF010144	Homo sapi
5	1301.4	90.2	1240001	9	HS886K2	Human DNA
6	1080.2	74.9	1381	6	AR051479	Sequence
7	1080.2	74.9	1381	6	AR072619	Sequence
8	1080.2	74.9	1381	6	AR073164	Sequence
9	592.2	41.1	160714	2	AC087434	Homo sapi
10	590.4	40.9	99370	9	AC005057	Homo sapi
11	587.8	40.8	65608	6	AX330738	Sequence
12	587.8	40.8	65608	6	AX335496	Sequence
13	587.8	40.8	65608	6	AX332242	Sequence
14	587.8	40.8	67046	9	U62293	Human LIM-k
15	587.8	40.8	148525	2	HSU63721	Human elast
16	587.8	40.8	121272	2	AC016675	Homo sapi
17	578.4	40.1	121272	2	AC004956	Homo sapi
18	567.6	39.4	170630	2	AC016905	Homo sapi
19	567.6	39.4	201300	2	AL137073	Human DNA
20	562.4	39.0	212956	2	AC090543	Homo sapi
21	558.2	38.7	129169	2	AC032038	Homo sapi
22	556.8	38.6	48478	2	AC026687	Homo sapi
23	556.8	38.6	137737	2	AC010378	Homo sapi
24	556.4	38.6	192330	2	AL590133	Homo sapi
25	555	38.5	91927	9	AC004771	Homo sapi
26	554.4	38.5	42565	9	AC011559	Homo sapi
27	551.8	38.3	211063	2	AC069254	Homo sapi
28	550.8	38.2	155949	9	AC091022	Homo sapi
29	550.8	38.2	156863	2	AC069249	Homo sapi
30	550.4	38.2	94121	9	HS7115D9	Human DNA
31	550.2	38.2	193897	2	AC069486	Homo sapi
32	548.2	38.0	174864	2	AC026125	Homo sapi
33	548.2	38.0	206052	2	AC010173	Homo sapi
34	545	37.8	173240	2	AC021586	Homo sapi
35	543.8	37.7	122146	9	AC011736	Homo sapi
36	541.2	37.5	62486	9	AL590093	Human DNA
37	540.2	37.5	171056	2	AC026318	Homo sapi
38	540.2	37.5	172248	2	AC012466	Homo sapi
39	538.2	37.3	129043	2	CN501DTD	Homo sapi
40	538.2	37.3	191946	9	CN501DTD	Homo sapi
41	537.2	37.3	110129	9	AC008956	Homo sapi
42	535.2	37.1	79528	9	HS466N1	Human DNA s
43	535.2	37.1	134896	9	AC006571	Homo sapi
44	535.2	37.1	150296	9	U91326	Human Chrom
45	534.8	37.1	112337	2	AC011509	Homo sapi

ALIGNMENTS

RESULT 1
LOCUS   AR051550   1442 bp   DNA   linear   PAT 29-SEP-1999
DEFINITION Sequence 120 from patent US 5830670.
ACCESSION   AR051550
VERSION   AR051550.1 GI:5974914
KEYWORDS   .
SOURCE   Unknown.
  ORGANISM Unknown.
            Unclassified.
REFERENCE   1 (bases 1 to 1442)
  AUTHORS   de la Monte,S. and Wands,J.R.
  TITLE   Neural thread protein gene expression and detection of Alzheimer's
            disease
  JOURNAL   Patent: US 5830670-A 120 03-NOV-1998;
FEATURES   Location/Qualifiers
     source   1..1442
                     /organism="unknown"
BASE COUNT   314 a   402 c   311 g   415 t
ORIGIN

Query Match   100.0%; Score 1442; DB 6; Length 1442;

Best Local Similarity 100.0%; Conservative 0; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;
Matches 1442;

```
Qy      1 ttttttttttttgagatggaagtttgcctctgttgttgcccaggctggagtggcaatggcgcaa 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ttttttttttttgagatggaagtttgcctctgttgttgcccaggctggagtggcaatggcgcaa 60

Qy     61 tctcagctcaccgcaacctccgcctcccgggttcaagcgattctcctgcctcagcctcc 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TCTCAGCTCACCGCAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCC 120

Qy    121 cagtagctgggattacaggcatgtgccaccacgctcggctaattttttgtattttag 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CAGTAGCTGGGATTACAGGCATGTGCCACCACGCTCGGCTAATTTTTTGTATTTTAG 180

Qy    181 tagagatggagtttctccatgttggtcaggctggtctcgaactcccgacctcaggtgatc 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TAGAGATGGAGTTTCTCCATGTTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGGTGATC 240

Qy    241 cctccgtcctccggccctcccaaagtgctgggattacaggcgtgagccaccgcgcccagtca 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CCTCCGTCCTCCGGCCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCGCCCAGTCA 300

Qy    301 tggctaattttttgtttggtgaaaacaggggtttcactgttgtccaagctggtctcctgagc 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TGGCTAATTTTTTGTTTGGTAGAAAACAGGGTTTCACTGTTGTCCAAGCTGGTCTCCTGAGC 360

Qy    361 tcaagcagtccacctgcctcagcctcccaaagtgctgggattacaggcgtgcagccgtgc 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TCAAGCAGTCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGCAGCCGTGC 420

Qy    421 ctggccttttatttttttattttttaagacacaggtgtccacctcttacccaggatgaag 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CTGGCCTTTTATTTTTTATTTTTTAAGACACAGGTGTCCCACTCTTACCCAGGATGAAG 480

Qy    481 tgcagtggtgtgatcacagctcactgcagcctcaacctcctgagatcaagcaatcctcctg 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TGCAGTGGTGTGATCACAGCTCACTGCAGCCTCAACCTCCTGAGATCAAGCAATCCTCCTG 540

Qy    541 cctcagcctcccaaagtagtctgggaccaaagacatgccaccactacacctggctaatttta 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CCTCAGCCTCCCAAAGTAGCTGGGACCAAAGACATGCACCACTACACCTGGCTAATTTTTA 600

Qy    601 tttttattttttaatttttttgagacagagtctcaacctctgtcacccaggctggagtgcagt 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 TTTTTATTTTTTAATTTTTTGAGACAGAGTCTCAACCTGTCACCCAGGCTGGAGTGCAGT 660

Qy    661 ggcgcaatctcggctcactgcaacctctgcctcccgggttcaagtgattattcctcctgccca 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GGCGCAAATCTTGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGTGATTATTCCTCCTGCCCA 720

Qy    721 gcctcctgagtagctgggactacaggcgcccacccagcctagctaatttttttgtatttt 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GCCTCCTGAGTAGCTGGGACTACAGGCGCCCACCACCAGCCTAGCTAATTTTTTGTATTTT 780

Qy    781 tagtagagatggggtttcacacatgttgtccaggttgatctctgaccttgtgatc 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 TAGTAGAGATGGGGTTCACACATGTTCGCCAGGTTGATCTCTGGACCTTGTGATC 840

Qy    841 tgcctgcctcggcctcccaaagtgctgtgggattacaggcgttgagccaccgcgcctta 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCACCCCGGCTTA 900

Qy    901 ttttaattttgtttgtttttgaaatggaagtcactctgttacccaggctggagtgcaat 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TTTTTAATTTTTTTGTTGTTTTTGAAATGGAAGTCACTCTGTTACCCAGGCTGGAGTGCAAT 960

Qy    961 ggcaaatctcggctcactgcaacctccgcctcccgggttcaaggagtgcaat 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GGCAAAATCTCGGCTCACTGCAACCTCCGCCTCCCGGGTTCAAGGAGTGCAAT 1020

Qy   1021 gcctcccaagcagcgtgggattacaggcaccgcacacacccgctaattttgtgtatttt 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 GCCTCCCAAGCAGCTGGGATTACAGGCACCGCACACACCCGCTAATTTTGTGTATTTT 1080
```

```
Db   1021 GCCTCCCAAGCAGCTGGGATTACAGGCGCCACCACCACCCGGCTTAATTTTTGTATTTT 1080

Qy   1081 tcattaggagcggggtttcacaccatattgtgcccaggctgttctccaaactcctgacctcaggt 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 TCATTAGAGGCGGGGTTTCACCATATTTGTGCCCAGGCTGGTCTCAAACTCCTGACCTCAGGT 1140

Qy   1141 gacccaccgcctcggcctcccaagtgctgggattacaggcgtgagcccacccag 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 GACCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCTCACCCAG 1200

Qy   1201 ccggctaatttttttgtatttttagtagagatgaggtctctgctatgttgcccaggct 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 CCGGCTAATTTTTTGTATTTTTAGTAGAGATGGGGTCTCTGCTATGTTGCCCAGGCT 1260

Qy   1261 ggtctcgaaacttctggcctcaagcaatcctccaaaatgagccacccagtca 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GGTCTCGAAACTTCTGGCCTCAAGCAATCCTCCAAAATGAGCCACCCAGTCA 1320

Qy   1321 catttttaaacagttacatctttattttagtactagaaagtaatacaataaaacatgt 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 CATTTTTAAACAGTTACATCTTTATTTTAGTACTAGAAAGTAATACAATAAAACATGT 1380

Qy   1381 caaacctgcaaattcagtagtaacagaagttctttttatttaacttttaaaacaaagctttag 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 CAAACCTGCAAATTCAGTAGTAACAGAAGTTCTTTTATTAACTTTTAAAACAAAGCTTTAG 1440

Qy   1441 ca 1442
          ||
Db   1441 CA 1442
```

RESULT 2
AR072690

LOCUS   AR072690   1442 bp   DNA   linear   PAT 28-AUG-2000
DEFINITION Sequence 120 from patent US 5948634.
ACCESSION   AR072690
VERSION   AR072690.1 GI:9999454
KEYWORDS   .
SOURCE   Unknown.
  ORGANISM Unknown.
            Unclassified.
REFERENCE   1 (bases 1 to 1442)
  AUTHORS   de la Monte,S. and Wands,J.R.
  TITLE   Neural thread protein gene expression and detection of alzheimer's
            disease
  JOURNAL   Patent: US 5948634-A 120 07-SEP-1999;
            Location/Qualifiers
FEATURES   Location/Qualifiers
     source   1..1442
                     /organism="unknown"
BASE COUNT   314 a   402 c   311 g   415 t
ORIGIN

Query Match   100.0%;   Score 1442;   DB 6;   Length 1442;
Best Local Similarity 100.0%;
Matches 1442; Conservative 0; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 ttttttttttttgagatggaagtttgcctctgttgttgcccaggctggagtggcaatggcgcaa 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TTTTTTTTTTTTGAGATGGAGTTTCGCTCTTGTTGCCCAGGCTGGAGTGCAATGGCGCAA 60

Qy     61 tctcagctcaccgcaacctccgcctcccgggttcaagcgattctcctgcctcagcctcc 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TCTCAGCTCACCGCAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCC 120

Qy    121 cagtagctgggattacaggcatgtgccaccacgctcggctaattttttgtattttag 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CAGTAGCTGGGATTACAGGCATGTGCCACCACGCTCGGCTAATTTTTTGTATTTTAG 180

Qy    181 tagagatggagtttctccatgttggtcaggctggtctcgaactcccgacctcaggtgatc 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TAGAGATGGAGTTTCTCCATGTTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGGTGATC 240
```

```
Qy   241 cctccgtctgggcctcccaaagtgctagtagatacaggactggccacactgccggccgtc  300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CCTCCGTCTGGGCCTCCCAAAGTGCTAGTAGATACAGGACTGGCCACCATGCCGGCTCGC  300

Qy   301 tggctaattttgtggtagaaacagggtttcactgattgtgcccaaagctggtctcctgagc  360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 TGGCTAATTTTGTTGGTAGAGAACAGGGTTTCACTGANTGTGCCCAAGCTGGTCTCCTGAGC  360

Qy   361 tcaagcagtcccaagtgctcagcctcccaaagtgctgggattacaggcgtgagccgtgc  420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 TCAAGCAGTCCCAAGTGCTCCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCGTGC  420

Qy   421 ctggccttcttttattttttttttaaggacacaggtgtccaactcttaccccaggatgaag  480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 CTGGCCTTTTTATTTTTATTTTTTTTAAGACACAGGTGTCCCACTCTTTACCCAGGATGAAG  480

Qy   481 tgcagtgtgtgatcaacagctcactgcagccttcaaactcctgagatcaagcatcctccctg  540
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 TGCAGTGGTGTGATCAACAGCTCACTGCAGCCTTCAAACTCCTGAGATCAAGCATCCTCCTG  540

Qy   541 cctccagcctcccaaagtagctgggactaacatgcacacactacacctggctaattttta  600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 CCTCAGCCTCCCAAGTAGCTGGGACCAAAGACATGCACACCACTACACCTGGCTAATTTTTA  600

Qy   601 ttttattttttaattttttgagacagagtctcaacctctgtcacccaggctggagtgcagt  660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 TTTTTATTTTTTAATTTTTTGAGACAGAGTCTCAACTCTGTCACCCAGGCTGGAGTGCAGT  660

Qy   661 gggcaaatcttggctcactgcaacctctgcctcccgggttcaagttattctcctgcccca  720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 GGGCAAATCTTGGCTCACTGCAACCTTCGCCTCCCGGGTTCAAGTTATTCTCCTGCCTCA  720

Qy   721 gcctcctgagtagctgggactacaggcgcccaccaccacgcccggctaattttttgtatttt  780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 GCCTCCTGAGTAGCTGGGACTACAGGCGCCCACCACCACGCCCGGCTAATTTTTTGTATTTT  780

Qy   781 tagtagagatggggtttcaccatgttccgccaggttgatctctgaacttgatc  840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 TAGTAGAGATGGGGTTTCACCATGTTCGGCAGGTTGATCTTGAATCTCTGGACCTTGTGATC  840

Qy   841 tgcctgcctcggcctcccaaagtgctgggattacaggcgtgagccaccacgcccggccta  900
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 TGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCACGCCCGGCCTTA  900

Qy   901 ttttaattttgttgtttgtttgtttttgaaatggaatctcactgttacccaggctgagtgcaat  960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TTTTTAATTTTGTTGTTTGTTTTTGAAATGGAATTCTCACTGTTGTTACCCAGGCTGGAGTGCAAT  960

Qy   961 ggccaaatctcgctcaacctctgcctcccgggcctcaaagcgattccctgctctca  1020
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 GGCCAAAATCTCGGCTCACTGCAACCTCTGCCTCCCGGGCTCAAGCGATTCTCCTGTCTCA  1020

Qy  1021 gcctcccaagcagctgggattacggggcaactcgacctgccacccaacctaattttttgtatt  1080
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021 GCCTCCCAAGCAGCTGGGATTACAGGCGTGTTTCACCANTTTGCCACACCCGCTCAATTTTTGTATT  1080

Qy  1081 tcattagaggcggggtttcaccatattgtggctcaaactcctgacctcaggt  1140
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 TCATTAGAGGCGGGGTTTCACCATATTTGTGGTCAAACTCCTGACCTCAGGT  1140

Qy  1141 gacccacctgcctcagcccttcccaaagtgctgctggggttattacaggcgtgagccacctgc  1200
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141 GACCCACCTGCCTCAGCCTTCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGTGC  1200

Qy  1201 ccggctaatttagataaaaaaaaatatgtagcaatgggggtcttgctatgttgcccaggct  1260
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201 CCGGCTAATTTAGATAAAAAAATATTGTAGCAATGGGGGTCTTGCTATGTTGCCCAGGCT  1260

Qy  1261 ggtctcaaacttctggcttcatgcaatcctccaaatgagccaacaacccagccagtca  1320
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261 GGTCTCAAACTTCTGGCTTCATGCAATCCTCCCAAATGAGCCACAACACCCAGCCAGTCA  1320

Qy  1321 cattttttaaacagttacatctttattttgtatactagaaagtaataacataaaacatgt  1380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

RESULT 3

```
LOCUS       AR073235                 1442 bp    DNA     linear   PAT 28-AUG-2000
DEFINITION  Sequence 120 from patent US 5948888.
ACCESSION   AR073235
VERSION     AR073235.1  GI:9999998
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1442)
  AUTHORS   de la Monte,S. and Wands,J.R.
  TITLE     Neural thread protein gene expression and detection of Alzheimer's
            disease
  JOURNAL   Patent: US 5948888-A 120 07-SEP-1999;
FEATURES             Location/Qualifiers
     source          1..1442
                     /organism="unknown"
BASE COUNT      314 a    402 c    311 g    415 t
ORIGIN
```

Query Match   100.0%; Score 1442; DB 6; Length 1442;
Best Local Similarity   100.0%; Pred. No. 0; Indels 0; Gaps 0;
Matches 1442;   Conservative 0; Mismatches 0;

```
Qy     1 ttttttttttgagatggagtcttgctctgtcgcccaggctggagttgcaatggcgcaa  60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 TTTTTTTTTTTGAGAGNGGAGTTTTCGCTCTTGTTGCCCAGGCTGGAGTGCAAT  60

Qy    61 tctcagctcacgcaaacctccgccctccgggttcaagcgattctcctgcctcagcctccc  120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 TCTCAGCTCACCGCAACCTCCGCCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCC  120

Qy   121 cagtagctgggattacaggcatgtgcaccacgcctcggctaattttttgtatttttttag  180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CAGTAGCTGGGATTACAGGCATGCGCCACCACGCCCGGCTAATTTTTGTATTTTTTTAG  180

Qy   181 tagagatggagttctccatgttggtcaggctggtctcgaactcctgacctcaggatgatc  240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 TAGAGATGGAGTTTCTCCATGTTGGTCAGGCTGGTCTCGAACTCCTGACCTCAGGATGATC  240

Qy   241 cctccgtctcggcctcccaaagtgctgggattacaggcgtgagccaccacgcccggcctgcc  300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CCTCCGTCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGTGCCCGGCCTGCC  300

Qy   301 tggctaattttgtgtgtagaaacaggggtttcactgatgtgcccaggctggtcctgagc  360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 TGGCTAATTTTGTGTGTTAGAAACAGGGTTTCACTGATGTGCCCAAGCTGGTCTCCTGAGC  360

Qy   361 tcaagcagtccaactgcctctgcctcccaaagtgctgggattacaggcgtgagccgccgtgc  420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 TCAAGCAGTCCAACTGCCTCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTCCCCAGCCGCCGTGC  420

Qy   421 ctggccttttttatttttatttttttaagacacaggtgtcccactcttaccccaggatgaag  480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 CTGGCCTTTTTATTTTTATTTTTTTTAAGACACAGGTGTCCCCACTCTTTACCCAGGATGAAG  480

Qy   481 tgcagtggtgtgatcaacagctcactgcagccttcaaactcctgagatcaagcatcctcctg  540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 TGCAGTGGTGTGATCACAGCTCACTGCAGCCTTCAAACTCCTGAGATCAAGCATCCTCCTG  540
```

```
Db  1321 CATTTTTTTAAACAGTTACATCTTTATTTTGTATACTAGAAAGTAATAACATAAAACATGT  1380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy  1381 caaaacctgcaaattcagtagtaacagagagttcttttatacttttataacttttaaacacaaagcttagag  1440
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381 CAAAACCTGCAAATTCAGTAGTAGTAAACAGAGAGTTCTTTTTATAACTTTTAAACAAAGCTTTAGAG  1440

Qy  1441 ca  1442
         ||
Db  1441 CA  1442
```

```
RESULT  4
LOCUS       AF010144      1442 bp    mRNA     linear   PRI 02-NOV-2001
DEFINITION  Homo sapiens neuronal thread protein AD7c-NTP mRNA, complete cds.
ACCESSION   AF010144
VERSION     AF010144.1  GI:3002526
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1442)
  AUTHORS   Monte,S.M., Ghanbari,K., Frey,W.H., Beheshti,I., Averback,P.,
            Hauser,S.L., Ghanbari,H.A. and Wands,J.R.
  TITLE     Characterization of the AD7c-NTP cDNA expression in Alzheimer's
            disease and measurement of a 41-kD protein in cerebrospinal fluid
  JOURNAL   J. Clin. Invest. 100 (12), 3093-3104 (1997)
  MEDLINE   98064067
  PUBMED    9399956
REFERENCE   2  (bases 1 to 1442)
  AUTHORS   de la Monte,S.M. and Wands,J.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-JUN-1997) MGH Cancer Center, Massachusetts General
            Hospital, 149 13th Street, Room 7308, Charlestown, MA 02129, USA
FEATURES             Location/Qualifiers
     source          1..1442
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /tissue_type="neuronal"
     CDS             15..1142
                     /note="contains expressed Alu-containing domains; 41 kDa
                     phosphoprotein; cell death and neuritic sprouting observed
                     in neuronal cells transfected with the AD7c-NTP cDNA"
                     /codon_start=1
                     /product="neuronal thread protein AD7c-NTP"
                     /protein_id="AAC08737.1"
                     /db_xref="GI:3002527"
                     /translation="MEFSLLLPRLECNGAISAHRNLRLPGSSDSPASSPVAGITGMC
                     THARLLIYFFLVEMEFLHVGQAGLELPTSDDPSVSASQSARYRYGHHARLCLANFCGR
                     NRVSLMCPSWSPELKQSTCLSLPKCWDYRRAAVPGLFLFFLRHRCPLLTQDDVQWCD
                     HSSLQPSTPEIKHPPASASQVAGHDMHHTWTWLLFIFIFNFLRQSLNSVYQAGVQWRN
                     LGSLQPLPPGFKFLFSCPSLLSSWDYRRPPRLANFVFVLVEMGFHHARLILISGPCDL
                     PASASQSAGITGVSHHARLIFNFCLFFEMESHSVTQAGVQWPNLGSLQPLPPGLKRFSC
                     LSLPSSWDYGHLPPHPANFCIFIRGGVSPYLSGWSQTPDLR"
BASE COUNT      314 a    402 c    311 g    415 t
ORIGIN


Query Match       100.0%;     Score 1442;     DB 9;     Length 1442;
Best Local Similarity   100.0%;
Matches 1442;     Conservative    0;     Mismatches 0;     Indels 0;     Gaps 0;

Qy      1 tttttttttttgaaatgggagttttttcgctcttgttcgccaggctggagtgcaatggcgcaa 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TTTTTTTTTTTTGAGATGGGAGTTTTTTCGCTCTTGTTCGCCAGGCTGGAGTGCAATGGCGCAA 60

Qy     61 tctcagctcaccgcaaactccgcgctccgggttcaagcgggattcctcgcctcagcctccc 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TCTCAGCTCACCGCAAACTCCGCGCTCCGGGTTCAAGCGGATTCCTCGCCTCAGCCTCCC 120

Qy    121 cagtagctggggattacagggcatgtgcaccacgctcgggtcaatttgtattttttttag 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CAGTAGCTGGGGATTACAGGCATGTGCACCACGCTCGGGTCAATTTGTATTTTTTTTAG 180

Qy    181 tagagatggagttctccacatgttggtcaggctggtctcgaactcccgacctcaggatgatc 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TAGAGATGGAGTTCTCCACATGTTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGATGATC 240

Qy    241 cctccgtctcggcctcccaaagtgctagatacagggcccaccatgcccggctctgcc 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CCTCCGTCTCGGCCTCCCAAAGTGCTAGATACAGGCGCCACCATGCCCGGCTCTGCC 300

Qy    301 tggctaattctttggtagaaaacaggggttcactgatgtgcccaagtcctggtctcctgagc 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TGGCTAATTTTTGTGGTAGAAAACAGGGGTTCACTGATGTGCCCAAGTCGGTGTCCTGAGC 360

Qy    361 tcaagcagtccacctgcctcagcctcccaaagtgctggggattacaggcgtgagccgtgc 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TCAAGCAGTCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCGTGC 420

Qy    541 cctcagccctcccaagtagctgggaccaaaagacatgcaacactacacctggctaatttttta 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CCTCAGCCCTCCCAAGTAGCTGGGACCAAAAGACATTGCAACACTACACCTGGCTAATTTTTA 600

Qy    601 tttttagttttttaatttttgagacagagtctcaactctgtcacccaggctggagtgcagt 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 TTTTTAGTTTTTTAATTTTTGAGACAGAGTCTCAACTCTGTCACCCAGGCTGGAGTGCAGT 660

Qy    661 ggcgccaatccttggctcactgcaacctctgcctcccgggttcaagttattctcctgcccca 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GGCGCCAATCTTGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGTTATTCTCCTGCCCCA 720

Qy    721 gcctcctgagtagctgggactacaggcgcccaccacacgcctagctaattttttttgtatttt 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GCCTCCTGAGTAGCTGGGACTACAGGCGCCCACCACACGCCTAGCTAATTTTTTTGTATTTT 780

Qy    781 tagtagagatggggtttcaccatgttgtccaggttgaccttgtgatctt 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 TAGTAGAGATGGGGTTTCACCATGTTGGCCAGGCTGGTCTCGATCTCCTGACCTTGTGATCC 840

Qy    841 tgcctgcctcggcctcccaaagtgctgtgggattacaggcgtgagccacccccgcgccta 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCCCGCCTA 900

Qy    901 ttttaatttttgttttgttttgaaatgagaaatccactctgttaccaggctggagtgcaat 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TTTTAATTTTTGTTTTGTTTTGAGAATGGAAATCCACTCTGTTACCAGGCTGGAGTGCAAT 960

Qy    961 ggccaaaatctcggctcactgcaacctctgcctcccgggctcaagcgattctccctgtctca 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GGCCAAAATCTCGGCTCACTGCAACCTCTGCCTCCCGGGCTCAAGCGATTCTCCCTGTCTCA 1020

Qy   1021 gcctcccaagcagctgggattacgggccacctgccacacaccccgctaattttgtattt 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 GCCTCCCAAGCAGCTGGGATTACGGGCCACCTGCCACACACCCCGCTAATTTTGTATTT 1080

Qy   1081 tcattagaggcggggtttcaccatatttgttcaaactcctgacctcaggt 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 TCATTAGAGGCGGGGTTTCACCATATTTGTTCAAACTCCTGACCTCAGGT 1140

Qy   1141 gacccacccctccctcagcccttccaaagtgctgggaacctccaccccag 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 GACCCACCGCCTCAGCCTTCCAAAGTGCTGGGAACCTCCACCGCCAG 1200

Qy   1201 ccggctaattagataaaaaatatgtagcaatgggcccctatggggtcttgctatgttgcccaggct 1260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 CCGGCTAATTAGATAAAAAATATGTAGCAATGGGCCCCTTATGGGGTCTTGCTATGTTGCCCAGGCT 1260

Qy   1261 ggtctcaaacttctggcttcatgcaatccttccaaatgagccacacaccagccagtca 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GGTCTCAAACTTCTGGCTTCATGCAATCCTTCCAAATGAGCCACAACACCAGCCAGTCA 1320

Qy   1321 catttttaaaacagttacaatctcttattttagtatactagaaaagtaataacataaacagt 1380
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 CATTTTTTAAAACAGTTACACTCTTTATTTTAGTATACTAGAAAGTAATACAATAAACACAGT 1380

Qy   1381 caaacctgcaaattcagtaagagagttcttttataacttttaaacaaagctttagg 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 CAAACCTGCAAATTCAGTAGTAACAGAGTTCTTTTATAACTTTTAAACAAAGCTTTAGAG 1440

Qy   1441 ca 1442
          ||
Db   1441 CA 1442
```

```
Qy    421 ctggcctttattatttattttttaaagacacaaggtgtccacactcttaccagatgaag 480
          |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db    421 CTGGCCTTTATTTATTTATTTTTTTAAGACACAAGGTGTCCCACTCTTACCCAGGATGAAG 480

Qy    481 tgcagtggtgtgatcacagctcactcgagccttcaactcctgagatcaagatcctcctg 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db    481 TGCACTGGTGTGATCACAGCTCACTCGAGCCTTCAACTCCTGAGATCAAGATCCTCCTG 540

Qy    541 cctcagcctcccaaagtagtgggaccaaagacatgcaccactacaccctggctaattta 600
          |||||||||||||| |||||||||||||||||||| |||||||||||||||||||||| ||
Db    541 CCTCAGCCTCCCAAGTAGCTGGGACCAAAGACATTGCACCACTACACCCTGGCTAATTTTTA 600

Qy    601 ttttattttttttttgagacagagtctcactctgtcacccaggctgagtgcagt 660
          |||| ||||| ||||||||||||||||||||||||||||||||||||||| |||||||
Db    601 TTTTTATTTTATTTTTTTTTGAGACAGAGTCTCACTCTGTCACCCAGGCTGGAGTGCAGT 660

Qy    661 ggcgcaatctcggctcactgcaacctctgcctcccgggttcaagttattctcctgccca 720
          ||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db    661 GGCGCAATCTCGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGTTATTCTCCTGCCCA 720

Qy    721 gcctcctgagtagctgggactacaggcgcccaacaccgcctagctaattttttgtatttt 780
          |||||||||||||||||||||||||||||||| ||||||||||||||||| |||||||||
Db    721 GCCTCCTGAGTAGCTGGGACTACAGGCGCCCACCACCGCCTAGCTAATTTTTGTATTTT 780

Qy    781 tagtagagatggggtttcaccatgttgccaggttgacctttgatctctggaccttgatc 840
          ||||||||||||||||||||||||||||||||||| |||| |||| |||||||||||||||
Db    781 TAGTAGAGATGGGGTTTCACCATGTTGCCAGGTTGATCTTGAACTTTGTGGACCTTGATC 840

Qy    841 tgcctgcctcggcctcccaaagtgctgggattacaggcgtgagccacacgcccggctta 900
          |||||||||||||||||||| ||||||||||||||||||||||||||| ||| ||||||||
Db    841 TGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCACCGCCCGGCTTA 900

Qy    901 tttttaattttgttttgttttgttttgaaatggaatctcaactctgttacccaggctgg 960
          ||||| |||||||| ||||||||||| ||||||||||||||||||||||||||||||||||
Db    901 TTTTTAATTTTGTTTTGTTTTGTTTTGAAATGGAATCTCAACTCTGTTACCCAGGCTGG 960

Qy    961 ggccaaatctcggctcactgcaacctctgcctccccgggttctcaagcgaattctcctgtctca 1020
          ||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db    961 GGCCAAATCTCGGCTCACTGCAACCTCTGCCTCCCGGGTTCTCAAGCGAATTCTCTGTCTCA 1020

Qy   1081 tcattagaggcgggggtttcaccatatttgtcaggctcaaactctgacctcaggt 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db   1081 TCATTAGAGGCGGGGTTTCACCATATTTGTCAGGCTCAAACTCCTGACCTCAGGT 1140

Qy   1141 gacccacctgcctcagcctcccaaagtgctgggattacgggcacctcaccag 1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 GACCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACGGGCACCTCACCAG 1200

Qy   1201 ccggctaatttttagatataaaaaatatgtagcaatgggggggtcttgctgttgcccaggct 1260
          ||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||| |
Db   1201 CCGGCTAATTTTAGATATAAAAAATATGTAGCAATGGGGGGGTCTTGCTGTTGCCCAGGCT 1260

Qy   1261 ggtctgaaactctggcctcaagtgatcctcccaaatgagccacaccaccgcagtca 1320
          ||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db   1261 GGTCTGAAACTCTGGCCTCATGCAATCCTTCCAAATGAGCCACACCACCGCAGTCA 1320

Qy   1321 cattttttaaacagttacatcttttattttagtatactagaaagtaataacaataaacatgt 1380
          |||||||||||||||| ||||||||||| ||||||||||||||||||||||||| ||||
Db   1321 CATTTTTTAAACAGTTACATCTTTTATTTTAGTATACTAGAAAGTAATACAATAAAACATGT 1380

Qy   1381 caaacctgcaaattcagtagtaacagagtctttattaaccttttataactttaaacaaagctttagag 1440
          |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db   1381 CAAACCTGCAAATTCAGTAGTAACAGAGTTCTTTATTAACTTTTATAAACTTTAAACAAAGCTTTAGAG 1440

Qy   1441 ca 1442
          ||
Db   1441 CA 1442
```

```
repeat_region   3385..3686
                /note="L1M4 repeat: matches 2200..2535 of consensus"
repeat_region   3709..4050
                /note="AluSx repeat: matches 1..310 of consensus"
repeat_region   4064..4212
                /note="THE1C repeat: matches 1..155 of consensus"
repeat_region   4358..4506
                /note="FRAM repeat: matches 1..371 of consensus"
repeat_region   4788..4878
                /note="MER5A repeat: matches 15..182 of consensus"
repeat_region   4879..5178
                /note="AluJo repeat: matches 54..149 of consensus"
repeat_region   5181..5484
                /note="AluSx repeat: matches 1..302 of consensus"
repeat_region   5485..5653
                /note="AluSg repeat: matches 1..305 of consensus"
repeat_region   5554..5788
                /note="AluJo repeat: matches 149..304 of consensus"
repeat_region   5789..6089
                /note="AluSx repeat: matches 1..132 of consensus"
repeat_region   6090..6260
                /note="AluSp repeat: matches 1..297 of consensus"
repeat_region   6282..6580
                /note="AluSx repeat: matches 132..312 of consensus"
repeat_region   6596..7224
                /note="AluY repeat: matches 1..296 of consensus"
repeat_region   7225..7293
                /note="L2 repeat: matches 1995..2748 of consensus"
repeat_region   7294..7753
                /note="AluJb repeat: matches 152..220 of consensus"
repeat_region   7756..8048
                /note="L2 repeat: matches 1543..1995 of consensus"
repeat_region   8050..8358
                /note="AluSg repeat: matches 1..279 of consensus"
repeat_region   8420..8987
                /note="AluY repeat: matches 1..309 of consensus"
repeat_region   8998..9090
                /note="L2 repeat: matches 873..1539 of consensus"
repeat_region   9060..9105
                /note="L2 repeat: matches 1543..1627 of consensus"
repeat_region   9180..9279
                /note="L2 repeat: matches 827..872 of consensus"
repeat_region   9489..9671
                /note="MIR repeat: matches 35..139 of consensus"
misc_feature    10046..10047
                /note="Tandem repeat. Region of tandem repeat each element
                48 base pairs and typical sequence:
                GAAGAACTGTGTTAGATTATTATATACATCCCTCTTGGATTC AGGTTCGT
                Restriction digest information (HINDIII) suggests 3 copies
                of the repeat are missing from this assembly."
repeat_region   10530..10594
                /note="L2 repeat: matches 2676..2750 of consensus"
repeat_region   10604..10678
                /note="MIR repeat: matches 12..87 of consensus"
repeat_region   11227..11408
                /note="MIR repeat: matches 20..232 of consensus"
repeat_region   11470..11776
                /note="AluSq repeat: matches 1..311 of consensus"
repeat_region   12244..12546
                /note="AluY repeat: matches 1..303 of consensus"
repeat_region   13021..13318
                /note="AluSx repeat: matches 1..295 of consensus"
repeat_region   13319..13389
                /note="MER5A repeat: matches 73..137 of consensus"
repeat_region   13390..13685
                /note="AluSq repeat: matches 1..297 of consensus"
repeat_region   13686..13730
                /note="MIR repeat: matches 27..73 of consensus"
repeat_region   13759..13823
                /note="MER5A repeat: matches 2146..2215 of consensus"
repeat_region   13824..13959
                /note="L2 repeat: matches 1..133 of consensus"
                /note="FLAM_A repeat: matches 1..133 of consensus"

repeat_region   13960..14115
                /note="MIR repeat: matches 2012..2146 of consensus"
repeat_region   14221..14479
                /note="L2 repeat: matches 2012..2146 of consensus"
repeat_region   14660..14709
                /note="MER21B repeat: matches 156..456 of consensus"
repeat_region   14953..15263
                /note="MIR repeat: matches 49..98 of consensus"
repeat_region   15508..15684
                /note="AluSx repeat: matches 1..311 of consensus"
repeat_region   15724..15857
                /note="MIR repeat: matches 2..218 of consensus"
repeat_region   15961..16086
                /note="FLAM_A repeat: matches 1..133 of consensus"
repeat_region   16087..16394
                /note="FAM repeat: matches 12..163 of consensus"
repeat_region   16395..16409
                /note="AluSg repeat: matches 3..310 of consensus"
repeat_region   16722..16804
                /note="FAM repeat: matches 163..174 of consensus"
repeat_region   16908..17483
                /note="L2 repeat: matches 6..114 of consensus"
repeat_region   17679..17881
                /note="L1MB3 repeat: matches 2582..2691 of consensus"
repeat_region   18516..18927
                /note="L1MB3 repeat: matches 5598..6185 of consensus"
repeat_region   18928..19043
                /note="L2 repeat: matches 2285..2490 of consensus"
repeat_region   19044..19149
                /note="MLT1D repeat: matches 114..505 of consensus"
repeat_region   19750..20088
                /note="FLAM_C repeat: matches 1..116 of consensus"
repeat_region   20092..20153
                /note="MLT1D repeat: matches 6..114 of consensus"
repeat_region   20170..20479
                /note="L2 repeat: matches 2273..2646 of consensus"
repeat_region   20499..20633
                /note="31 copies 2 mer aa 81% conserved"
repeat_region   20639..20946
                /note="AluSx repeat: matches 3..312 of consensus"
repeat_region   20959..21232
                /note="L2 repeat: matches 2623..2750 of consensus"
repeat_region   21240..21532
                /note="AluJb repeat: matches 1..308 of consensus"
repeat_region   21534..21837
                /note="AluSg repeat: matches 1..297 of consensus"
repeat_region   21839..21936
                /note="AluSp repeat: matches 1..293 of consensus"
repeat_region   21952..22283
                /note="MER5B repeat: matches 2..310 of consensus"
repeat_region   22303..22605
                /note="L2 repeat: matches 84..175 of consensus"
repeat_region   22665..22969
                /note="AluY repeat: matches 1811..2198 of consensus"
repeat_region   23308..23605
                /note="AluSg repeat: matches 1..301 of consensus"
repeat_region
                /note="AluSp repeat: matches 1..304 of consensus"
repeat_region
                /note="AluSp repeat: matches 1..298 of consensus"
```

Query Match   Best Local Similarity   90.28%;   Conservative   97.48%;   Score 1301.4;
Matches 1409;   Conservative   0;   Mismatches   26;   Pred. No. 0;   DB 9;   Indels   12;   Length 124001;   Gaps   8;

```
Qy      1 tttttttttgagatggagtttcgctcttgttgcccaggctggagtgcaatggcgcaa   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6086 TTTTTTTTTTTTGAGATGGAGTTTCGCTCTTGTTGCCCAGGCTGGAGTGCAATGGCGCAA 6027

Qy     61 tctcagctcacgcaacctccgcctcccgggttcaagcgattctcctgcctcagcctcc  120
          ||||||||||| |||||||||||||||||||||||||||||||||||||||||||||| |
Db   6026 TCTCAGCTCACCGCAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCC 5967

Qy    121 cagtagctgggattacaggcgctcggcctaattttgtatttttttag            180
          ||||||||||||||||||| |||||                     |||||||||||||||
Db   5966 TAGTAGCTGGGATTACAGGCGCATGTGCCACCACGCCTCGGCTAATTTTTGTATTTTTTTAG 5907
```

```
RESULT   6
LOCUS       AR051479                1381 bp    DNA        linear   PAT 29-SEP-1999
DEFINITION  Sequence 49 from patent US 5830670.
ACCESSION   AR051479
VERSION     AR051479.1  GI:5974843
KEYWORDS
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1381)
  AUTHORS   de la Monte,S. and Wands,J.R.
  TITLE     Neural thread protein gene expression and detection of Alzheimer's
            disease
  JOURNAL   Patent: US 5830670-A 49 03-NOV-1998;
FEATURES             Location/Qualifiers
     source          1..1381
                     /organism="unknown"
BASE COUNT      296 a      384 c      302 g      399 t
ORIGIN

Query Match      74.9%;  Score 1080.2;  DB 6;  Length 1381;
Best Local Similarity   94.4%;   Pred. No 1.6e-311;
Matches 1317;  Conservative   0;  Mismatches 53;  Indels 25;  Gaps 18;
```

```
Qy      2 ttttttttttgagatggagtttcgtcttgttgcccaggctggagtgcaatggcgcaat   61
          | |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TTTTTTTTTTGAGATGGAGTTTCGCTCTTGTTGTCCCAGGCTGGAGTGCAANGGCGCAAT   60

Qy     62 ctcaagtcaacgcaacctccgcctcccggttcaagcgattctcctgcctcagcctccc  121
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CTCAGCTCACCGCAACCTCCGCCTCCCGGGTTCAAGCGANTTCTCCTGCCTCAGCCTCCC  120

Qy    122 agtagctgggattacaggcatgtgcaccacgcggcctaattttgtatttttttttagt  181
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 AGTAGCTGGGATTACAGGCANGTGCA-CCAGCCTCGGCTAATTTTGTATTTTTTTTTAGT  179

Qy    182 agagatggaggttt--cccatgttggtcaaggctggtctcgaactccccgaacctcag  239
          |||||||||||||       ||||||||||||||||||||||||||||||||||||
Db    121 AGAGANGGAGTTTAACTCCATGTTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGATGAT  239

Qy    240 ccctccgtctcggccctccaaagtgct---agatacaggactggccaccatggccggct  295
          |||||||||||||||||||||||||       |||||||||||||||||||||||||
Db    240 CTCCCGTCTCGGCCCTCCAAAGTGCTGGGATTACAGGCGCCACCATGGCCGGCCT       299

Qy    296 ctgagctcaagcagtccaactcctcagcctcccaaagtgctgggattacgggtgtcag  355
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    300 CTGAGCTCAAGGAGTCCAACTCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGT-CAG  359

Qy    356 gcttag 1438
          | |||
Db   4649 GTATCAG 4643
```

```
Qy    181 tagagatgggagtttctccaccatgttggtcaggctgtctcgaacctcccgacctcagatgatc  240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5906 TAGAGAtGGAGttTtTCCATGTTGGTCAGGCTGGTCTCGAACCTCCCGACCTCAGATGAtC  5847

Qy    241 ccctccgtctccggcctcccaaagtgct---agatacaggactggccaccatggccggcctc  296
          ||||||||||||||||||       ||||||||||||||||||||||||||||
Db   5846 CTCCCGTCTCGGGCCTCCAAAGTGCTGAGATTACAGGCATGAGCCACCACGCCCGGCCTC  5787

Qy    297 tgcctggcttaattttgtggttagaaacaggggtttcactgatgttgcccaagctggtctcc  355
          |||||||||||||||||||||||||||||||||||||||||| |||||||||
Db   5786 TGCCTGGCTTAATTTTGTGGTAGAAACAGGGGTTTCACTGATGTTGCCCAAGCTGGTCTCC  5727

Qy    356 tgagctcaagcagtccacctgcttcagcctcccaaagtgctggcttcactgttaccaggcgt--gca  413
          ||||||||||||||||||||||||||||||||||||||||||||        ||
Db   5726 TGAGCTCAAGCAGTCCACCTGCTTCCCAAAGTGCTCGGATTACAGGCGTGAGCC       4643

Qy    414 gccgtgcctggcctttattttattttttttaagacacaggtgtcccacttacccag  473
          |||||||||||||||||||||||||||||||||||||||| |||
Db   5666 ACCGTGCCTGGCCTTTTATTTTATTTTTTTTAAGACACAGGTGTCCCACTCTTACCCAG  5607

Qy    474 gatgaagtgcagtgtgtgatcaacagctcactgcagccttcaactcctgagatcaagca  532
          ||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db   5606 GATGAAGTGCAGTGGTGTGATCAACAGCTCACTGCAGCCTTCAACTCCTGAGATCAAGCAA  5547

Qy    593 aattttattttattttttttgagacagagtctc-actctgtcacccaggctgg  652
          ||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db   5546 AATTTTATTTTTATTTTTTTTGAGACAGAGTCTC-ACTCTGTCACCCAGGCTGG  5487

Qy    653 agtgcagtggcgcaatcttggctcactgcaacctctgcctcccgggttcaagtgattctc  712
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5486 AGTGCAGTGGCGGCAANTCTTGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGTTATTCTC  5428

Qy    713 ctgcccaagctcctgagtagctgggactacaggcgcgccaccaccaggctagctaatttttt  772
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5427 CTGCCCAAGGCTTCCTGAGTAGCTGGGACTACAGGCGCGCCACCACCACGCCCGGGCTAATTTTTT  5368

Qy    773 tgtattttttagtagagatgggg--ttcacacatgttcgccagggttgatctctggac  831
          ||||||||||||||||||||       |||||||||||||||||||||||||
Db   5367 CTGCCCCAAGGGGCTTCAGAGTAGCTGGGACTACAGGGGTTTCACCATGGTTGGTCTTTGGTTGAGTCTCTTGAC  5248

Qy    832 cttgtgatctgcctgcctcggccctccaaagtgctcggattacaggcgtgagccaccacg  891
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5307 TGTATTTTTAGTAGAGANGGGGGTTTCACCATGGTTCCTCCTGAGTTGGTCTTTGAGTCTCTTTGAC  5248

Qy    892 cccggcttatttttaatttttgttttgttgaaatgaatccacttcactctgttaccaggctg  951
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5247 CTTGTGANCTGCCTGCCTCGGCCCTCCAAAGTGCTGGGATTACAGGCGTGAGCCACCACG  5188

Qy    952 gagtgcaaatgcgcaaaatcggctcaggctacgccaacctctgcctcccgggcttcaaagt  1011
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5187 CCCGGCCTATTTTTAATTTTGTTTTGTTGTTTGAAATGAATCCACTCACTCTGTTACCAGGCTG  5128

Qy   1012 cctgtctcaagcctccaagcagctgggattacgggcacctgcaccacaccccgctaatt  1071
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5127 GAGTGCAATGGCACAATCTCGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGCGANTTCT  5068

Qy   1072 tttgtatttttcattagaggcgggggttcaccatacagggtctcaaactcctg  1131
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5067 CCTGTCTCAGCCTCCCAAGCAGCTGGGATTACGGGCACCTGCCA-ACACCCCGCTAATT  5010

Qy   1132 acctcaggtgaccacctgcctcagcctcccaaagtgctgggattacaggcgtgagccac  1191
          |||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db   5009 TTTGTATTTTTCATTAGAGGCGGGGGGTTCAGGCGGTGAGCCAC  4950

Qy   1192 ctcaccacacggctaattttagataaaaaaatatgtagcaatggggggttctgtcttatgtt  1251
          |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db   4949 ACCTCAGGTGACCACCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCAC  4890

Qy   1252 gcccaagctggtctcaaacttctggcttcatgtgcaatccttccaaatgaagccacacaccc  1311
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4889 CTCACCACCACGGCCTAATTTTTAGATAAAAAAATATGTAGCAATGGGGGGTTCTTGCTANGTT  4830

Qy   1252 gcccaagctggtctcaaacttctggcttcatgtgcaatccttccaaatgagccacacaccc  1311
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4829 GCCCAAGGCTGGTCTCGAAACTTCTGGCTTCAAGTGCAATCCTTCCAAATGAGCCACACACCC  4770

Db   4769 AGCCAGTCACATTTTTAAACAGTTACACTTTANTTTTAGTATACTAGAAAGTAATTACAA  4710

Qy   1312 agccagtcacatttttaaacagttacatctttattttagtatactaagaaagtaataacaa  1371
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4769 AGCCAGTCACATTTTTAAACAGTTACACTTTANTTTTAGTATACTAGAAAGTAATTACAA  4710

Qy   1372 taaacatgttcaaactgcaaattcagtagtaacagagttcttttatcatttttaaacaaa  1431
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4709 TAAACANTGTCAAACCTGCAAATTCAGTAGTAACAGAGTTCTTTTTATAACTTTTTAAACAAA  4650

Qy   1432 gctttag 1438
          |  |||
Db   4649 GTATCAG 4643
```

```
RESULT      7
LOCUS       AR072619                1381 bp    DNA     linear   PAT 28-AUG-2000
DEFINITION  Sequence 49 from patent US 5948634.
ACCESSION   AR072619
VERSION     AR072619.1  GI:999383
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1381)
  AUTHORS   de la Monte,S. and Wands,J.R.
  TITLE     Neural thread protein gene expression and detection of alzheimer's
            disease
  JOURNAL   Patent: US 5948634-A 49 07-SEP-1999;
FEATURES             Location/Qualifiers
     source          1..1381
                     /organism="unknown"
BASE COUNT      296 a    384 c    302 g    399 t
ORIGIN

Query Match       74.9%;  Score 1080.2;  DB 6;  Length 1381;
Best Local Similarity   94.4%;   Pred. No. 1.6e-311;
Matches 1317;  Conservative   0;  Mismatches   53;  Indels   25;  Gaps   18;
```

```
Qy      2  tttttttttgagagtggagtttttcgctcttgttgcccaggctggagtgcaatggcgcaat   61
           ||||||||| ||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db      1  tTTTTTTTTTTGAGAtGGAGTTTTTCGCTCTTGTTGCCCAGGCTGGAGTGCAATGGCGCAAT   60

Qy     62  ctcagctcacgcagcaacctccgcctcccgggttcaagcgattctcctgcctcagcctccc  121
           ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db     61  CTCAGCTCACGCAGCAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCC  121

Qy    122  agtagctgggattacaggcatgtgccaccacgctcgggctaattttgtatttttagt      181
           ||||||||||||||||||||| |||||| |||||||||||||| ||||||| ||||||||||
Db    121  AGTAGCTGGGATTACAGGCATGtGCCA-CCACGCTCGGGCTAATTTTTGTATTTTTTTAGT  181

Qy    182  agaatgggagtgtt--ctccatgtgtggtcaaggctggtctcggaactccgaacctcaagatgat  239
           ||||| |||||||||||  |||||||||||| |||||||||||||||||||||||||||||
Db    180  AGAGAtGGGAGTTtAACTCCATGtTGGTCGAGGCTGGTGTCTCGGAACTCCGAACCTCAGATGAT  239

Qy    240  ccctccgtctcgggcctcccaaagtgct---agatacaggactggccaccatgcccgg-ct  295
           ||||||||||||||||||| |||||||||   |||||||||||||| ||||||||||||| ||
Db    240  CTCCCGTCTCGGGCCTCCCAAAGTGCTGGGATTACAGGCATGCCCACCACATGCCCGGCCT  295

Qy    296  ctgcctggctaattttttggtagaaacacaggttcactgatg-tgcccaagctggtctc    354
           ||||||||||||||||||| || ||||| |||||||||||| ||||||||||||||||||||
Db    300  CTGCCTGGCTAATTTTGTGGTAGAAACAGGGTTCACTGAtGTtGCCCAAGCTGGTCTC     359

Qy    355  ctgagctcaagcagtccaacctgcctcagccctcccaaagtgctgggattacaggcgtgcag  414
           ||||||||||||||||||||| ||||||||||||||||||||||||||||||||| ||||||
Db    360  CTGAGCTCAAGCAGtCCAACGtCCACCTGCCTCAGCCCTCCCAAAGtGCTGGGATTACAGGCGTt-CAG  418

Qy    415  ccgtgcctggcctttcttatttttttagtacacaaggttgtccaactcttacccagg    474
           ||||||||||| |||||||||||||||||||||||||| ||||||||||||||||||||
Db    419  CCGTGCCTGGCCTttttTATTTTTTTAGTACACAAGGTtGTCCAACTCTTACCCAGG  474

Qy    475  atggaagtgcagtgtgtgatcacagcctcactgcagcctccaactctgagatcaagc-at  533
           |||||||||||||||||||||||||||||||||||||||| ||||| |||||||||||||
Db    479  ATGGAAGTGCAGTGtGTGATCACAGCTCACTGCAGCCTTCAACTCTGAGATCAAGCAAT  533

Qy    534  cctcctgcctccagaagtagctgggactacaggcgccacccacgccctagctaatttttttt   593
           ||||||||||||||||| |||| ||||||||||||||||||||||||||||||||||||||||
Db    539  CCTCCTGCCTCAGCCTCCAAGtAGCTGGGACCACAACACTACACCTACACCGGCta  593

Qy    594  atttttatttttattttttgagacagagtctcaactctgtcacccaggctgga    653
           |||||||||| ||||||| |||||||||||||||||||||||||||||||||||
Db    598  ATTTTTATTTTTATTTTTTTGAGACAGAGTCTC-ACTCTGTCACCCAGGCTGGA  653

Qy    654  gtgcagtggcgcaatctctggctcactgcaacctctgcctcccgggttcaagtgattatctcc  713
           |||||||||||||||||||||| |||||| ||||||||||||||||||||||||||||||||||
Db    657  GTGCAGTGGCGCAATCTCGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGTTGATTATTCTCC  713

Qy    714  tgcccagcctcccctgagctgggactacaggtgcccacccacccacccggctaatttttttt   773
           |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db    717  TGCCCAGCCTCCTGAGTACCTGGGACTACAGGCGCCCACCACGCCCACCACGGCTAATTTTTTT  773

Qy    774  gtattttagtagagatgggg-ttcaccatgttcgccaggttgatctggactct      832
           ||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db    774  GTATTTTAGTAGAGAtGGGGTTTCACCATGtTCGCCAGGGTTCaagttaatttttt     832
```

(left column continuation)

```
Qy    475  atgaagtgcagtgtgtgatcacagctcactgcagcctcctgagatcaagc-at     533
           ||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db    479  ATGAAGTGCAGTGGTGTGATCACAGCTCACTGCAACCTCTGCCTCTGAGATCAAGCAAT  538

Qy    534  cctcctgcctcagtagctgggactacaggcaacatgcaccactacacctggcta     593
           |||||||||||| ||||||||||||||||||  ||||||||||||||||||||||
Db    539  CCTCCTGCCTCAAGtAGCTGGGACCACAACAtGCACCACTACACCTACACCTGG-TA  597

Qy    594  attttatttttatttttttgagacagagtctcaactctgtcacccaggctgga     653
           |||||| |||||||||||||||||||||||| ||||||||||||||||||||||
Db    598  ATTTTTATTTTATTTTTTTAATTTTTTTAATTTTTTTGAGACAGAGTCTC-ACTCTGTCACCCAGGCTGGA  656

Qy    654  gtgcagtggcgcaatctctggctcactgcaacctctgcctcccgggttcaagttattctcc     713
           |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db    657  GTGCAGTGGCGCAATCTCGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGTTATTATTCTCC  713

Qy    714  tgccccagcctcctgagtagctgggactacaggcgcccacccacgccctagctaatttttt      773
           ||||| ||||||||||||||||||||||||||||| ||||||||||||||| ||||||||
Db    717  TGCCCAGCCTCCTGAGTACCTGGGACTACAGGCGCCCACCACGCCCACCACGGCTAATTTTTTTT  773

Qy    774  gtattttagtagagatgggg-ttcaccatgttcgccaggttcaactctttttt     832
           |||||||||||||||||||| |||||||||||||||||||| ||||||||||||
Db    777  GTATTTTAGTAGAGAtGGGGGTTTCACCATGtTGGCCAGGCTGGTCTC          832

Qy    833  ttgtgatcctgcctgcctcggcctcccaaagtgctgggattacaggcgtgagccacacg   892
           | |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db    837  TTGTGATCTGCCTGCCTCGGCCTCCCAAAGtGCTGGGATTACAG--GtCGtGACTCCAC  894

Qy    892  cccggcttatttttaattttttgttgtttgaaatggaatctcactctgttacccaggctg   951
           |||||||||||| |||||||||| ||| |||||||||||||||||||||||||||||||||
Db    895  GCCGGCCtATTTTTTAATTTTGTTGTTTTTGTTTGAAATGGAATCTCACTCTGTTACCCAGGCTG  951

Qy    952  gagtgcaatggcaaaatctcggctcactgcaacctctgcctcccgggctccaagcgattct  1011
           ||||||||||||||||||||||||||||||||||||| |||||||||||||||||| |||||
Db    955  GAGTGCAATGG-CAAATCTCGGCTACTGCAACCTCCGCCTCCCGGGCTCCCAAGCGATTCT  1012

Qy   1012  cctgtctcagcctcccaagcagcagctgggattacgggcgccacacacccccgctaatt  1071
           |||| |||||||||||||||||||||| |||||||| ||||||||||||||||| |||||
Db   1013  CCTGTCTCAGCCTCCCAAGCAGCTGGGATTACGGG---ACCTGCCACCACACCCGGCTAATT  1071

Qy   1072  tttgtattttttattagagacggggttctcaccatgttggtcaggct-ggtctcaaactcct  1130
           ||| ||||||| ||||||||||||||||||||||||||||| ||||| |||||||||||||||
Db   1072  TTTGTATTTTATTTAGAGGCGGGGTTTCAttAACCATAtTtGTCAGGCTGGTCTCAAACTCCT  1128

Qy   1131  gacctcaggtgacccacctgcctcagcctcccaaagtgctgctggattacaggcgtgagcca  1190
           |||||||| ||||||||||||| ||||| |||||||||| |||||||||||||||||||||||
Db   1129  GACCTCAGGTGACCCACCTGCCTCAGCCTCCCAAAGtGCTGGGATTACAGGCGTGAGCCA  1190

Qy   1191  cctcaccccagccgcctattttagataaaaaaataatgtgagcaatggggggtcttgctatgt  1250
           |||||||||||||||||||||| |||||||||| ||| |||||||||||| |||||||||||||
Db   1189  CCTCACCCAGCCGGCTAATTTGtgaatAAAAAAATAtGtGAGCAAtGGGGGGtCttgctatgt  1250

Qy   1251  tgcccaggctggtctctcaaacttctggctcatgtgcaaaatgagccacacacacc  1310
           |||||||||||||||||| |||||||||||| ||||||||||||||||||||||||
Db   1247  TGCCCAGGCTGGTCTCCAAACTTCTGGCTCAAtGTGCAAAATGGCCACACC  1306

Qy   1311  cagccagtcacacattttaaacagttacatcttattttaagtactagagaaagtaaataca  1370
           |||||||||||| ||||| ||||||||| |||||||||||||||||||||||||||||||
Db   1307  CAGCCAGTCACATTTTTAAAACAGTTACATCTTATTTTAGTATACTAGAGAAAGTAAATACA  1366

Qy   1371  ataaaacatgtcaaac  1385
           ||||| ||| ||||||
Db   1367  ATAAAACATGTCAAAC  1381
```

```
Db   777 GTATTTTTAGTAGAGATGGGGTTTCACCATGTTCGCCAGGTTGGTCTTGATCTCTTGACC 836
Qy   833 ttgtgatctgctggcctcg-gggcct-cccaaagtgctgggattacaggcgtgagccaccacg 891
Db   837 TTGTGATCTGCTGGCCTCGGGGCCTCCCAAAGTGCTGGGATTACAG--GTCGTGACTTACAG 894
Qy   892 cccggccttattttgttttgttttgttttgaaatgctcactctgttaccaggctg 951
Db   895 GCCCGGCCCTATTTTTAATTTTTGTTTTGTTTTGAAATGGAATCTCACTCTGTTACCCAGGGTCG 954
Qy   952 gagtgcaatggcgcaaatctcggctcactgcaactctcgcctcccgggttcaaggggattct 1011
Db   955 GAGTGCAATGG-CAAATCTCGGCTACTCGCCAACCTTCGCCTCCCGGG-TCAAGCGATTCT 1012
Qy  1012 cctgtctcagcctcccaagcagctggggattacaggcgcaccgcacaccccgctaatt 1071
Db  1013 CCTGTCTCAGCCTCCCAAGCAGCTGGGATTACGGG--ACCTGCACCACCACCCCGCCTAATT 1070
Qy  1072 tttgtattttcattagagacggggtttcaccacatatttgtcaggct-ggtctcaaactcct 1130
Db  1071 TTTGTATTTTCATTAGAGACGGGGTTT--ACCATATTTGTCAGGGTCTCAAACTCCT 1128
Qy  1131 gacctcaggtgacccacccgcctcagcctcccaaagtgctgggattacaggcgtgagcca 1190
Db  1129 GACCTCAGGTGACCCACCCGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCA 1188
Qy  1191 cctcacccagccggctaatttagtagagacggggtttcaccatgttggccaggctggtct 1250
Db  1189 CCTCACCCAGCCGGCTAATTTTGTGGAATAAAAAAATATGTAGCAATGGGGG--TCTGCTATGT 1246
Qy  1251 tgcccaggctggtctccaaacttctggcttcatgcaatcctcccaaagtgagccacacacc 1310
Db  1247 TGCCCAGGCTGGTCTCAAACTTCTGGGCTTCAGTCAATCCTCC-----TA 1306
Qy  1311 cagccagtgcaacatttttaaacagttacatctttattttttagtactagaagtaataca 1370
Db  1307 CAGCCAGTACACATTTTTTTAAACAGTTACATCTTTATTTTTTAGTATACTAGAAGTAATTACA 1366
Qy  1371 ataaacatgtcaaac 1385
Db  1367 ATAAACATGTCAAAC 1381
```

RESULT 8
AR073164

```
LOCUS       AR073164              1381 bp    DNA        linear    PAT 28-AUG-2000
DEFINITION  Sequence 49 from patent US 5948888.
ACCESSION   AR073164
VERSION     AR073164.1  GI:9999927
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1381)
  AUTHORS   de la Monte,S. and Wands,J.R.
  TITLE     Neural thread protein gene expression and detection of Alzheimer's
            disease
  JOURNAL   Patent: US 5948888-A 49 07-SEP-1999;
FEATURES            Location/Qualifiers
     source         1..1381
                    /organism="unknown"
BASE COUNT     296 a    384 c    302 g    399 t
ORIGIN
```

Query Match 74.9%; Score 1080.2; DB 6; Length 1381;
Best Local Similarity 94.4%; Pred. No. 1.6e-311;
Matches 1317; Conservative 0; Mismatches 53; Indels 25; Gaps 18;

```
Qy    2 ttttttttttgagatggagtttcgctcttgttcgcctcggctggagtgcaatggcgccaat 61
Db    1 TTTTTTTTTTTTGAGATGGAGTTTCGCTCTTGTTGCCCAGGCTGGAGTGCAATGGCGCCAAT 60
```

```
Qy   62 ctcagctcaccgcaacctccgcctcccgggttcaaggggattctcctgcctcagcctccc 121
Db   61 CTCAGCTCACCGCAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCCC 120
Qy  122 agtagctgggattacaggcgtgcaccaccacgctcggcgctaatttttgtattttttagt 181
Db  121 AGTAGCTGGGATTACAGGCATGTGCCACCACGCTCGGCTAATTTTTGTATTTTTTAGT 179
Qy  182 agagatggagtttt--ctccatgtcggtcaggctgtctcgaactcccgacctcaggtgat 239
Db  180 AGAGATGGGAGTTTAACTCCATGTTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGGTGAT 239
Qy  240 ccctccgtctccggccctcccaaagtgct--agatacaggactggccaccatgcccggct-ct 295
Db  240 CTCCCGATCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCACGCCCGGCCT 299
Qy  296 ctgagctcaagcagtccaactccgcctcccaaagtgctgggattacaggcgtgtctc 354
Db  300 CTGCCTGGCTCAATTTCTGTAGAAACAGGGTTTCACTGTGTTGCCCAGGCTGTCTC 359
Qy  355 cctgagctcaagcagtccaactccgcctcccaaagtgctgggattacaggcgtg-at 414
Db  360 CTGAGGTCAAGCAGTCCAACTCCGCCTCCCAAAGTGCTGGGATTACAGGCGTT-CAG 418
Qy  415 ccgtgcctggccttttttttattttttttttataagacacagggtctcccactcaccagg 474
Db  419 CCGTGCCTGGCCTTTTTTTATTTATTTTTTTTTAAGACACAGGGTGTCACCTCTTACCCAGG 478
Qy  475 atgaagtgcagtggtgtgatcacagctcactgcaactccacctccgaggtcaagc-at 533
Db  479 ATGAAGTGCAGTGGTGTGATCTCAGCTCACTGCAACCTCCGAGGTTCACCTGATTCTCC 538
Qy  534 cctccctgcctcagcctcccaagtagctgggaccaaagacagtgcaccaccactacacc 593
Db  539 CCTCCCTGCCTCAGCCTCCCAAGTAGCTGGGACCACAACATGCACCACCACCACCTGAGG-TA 597
Qy  594 attttttattttttttttttgagacagagtctcaactctgtcacccaggctgga 653
Db  598 AATTTTTATTTTTTTGAGACAGAGTCTC-ACTCTGTCACCCAGGCTGGA 656
Qy  654 gtgcaagtggcgcaatctcggctcactgcaacctctgcctcccgggttcaagtgattctcc 713
Db  657 GTGCAGTGGCGCAATCTCGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGTTTATTCTCC 716
Qy  714 tgcccagcctcctgagtagctgggactacaggcgcccaacaccacgcctagctaatttttt 773
Db  717 TGCCCCAGCCTCCTGAGTAGCTGGGACTACAGGCGCCCACCACCACGCGCCGGCTAATTTTTT 776
Qy  774 gtattttttagtagagatgggg--ttcaccatgttcggccaggttgatctctggacc 832
Db  777 GTATTTTTAGTAGAGATGGGGTTTCACCATGTTCGCCAGGTTGGTCTTGATCTCTTGACC 836
Qy  833 ttgtgatctgctggcctcg-gggcct-cccaaagtgctgggattacaggcgtgagccaccacg 891
Db  837 TTGTGATCTGCTGGCCTCGGGGCCTCCCAAAGTGCTGGGATTACAG--GTCGTGACTTACAG 894
Qy  892 cccggccttattttgttttgttttgttttgaaatgctcactctgttaccaggctg 951
Db  895 GCCCGGCCCTATTTTTAATTTTTGTTTTGTTTTGAAATGGAATCTCACTCTGTTACCCAGGGTCG 954
Qy  952 gagtgcaatggcgcaaatctcggctcactgcaactctcgcctcccgggttcaaggggattct 1011
Db  955 GAGTGCAATGG-CAAATCTCGGCTACTCGCCAACCTTCGCCTCCCGGG-TCAAGCGATTCT 1012
Qy  1012 cctgtctcagcctcccaagcagctggggattacaggcgcaccgcacaccccgctaatt 1071
Db  1013 CCTGTCTCAGCCTCCCAAGCAGCTGGGATTACGGG--ACCTGCACCACCACCCCGCCTAATT 1070
Qy  1072 tttgtattttcattagagacggggtttcaccacatatttgtcaggct-ggtctcaaactcct 1130
Db  1071 TTTGTATTTTCATTAGAGACGGGGTTT--ACCATATTTGTCAGGGTCTCAAACTCCT 1128
Qy  1131 gacctcaggtgacccacccgcctcagcctcccaaagtgctgggattacaggcgtgagcca 1190
```

```
Db   1129 GACCTCAGGTGACCCACCTTGCCTCAGCCTTCCAAACTTCTGGGATTACAGGCGTGAGCCA 1188
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1191 cctcaccaccggggtcaatttagataaaaaatatgtagacaatggggggtcttgctatgt 1250
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1189 CCTCACCCAGCCGGGTCAATTTGGAATAAAAAATAATGTAGCAATGGGGG--TCTGCTATGT 1246

Qy   1251 tgcccaggcgctggttccaaacttctggcttcatgcagccacacacacc 1310
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1247 TGCCCAGGCTGGTCTCAAACTTCTGGCTTCAGTCAATCCTTCCAAATGAGCCACACACACC 1306

Qy   1311 cagccagtcacatttttaaacagttacatctttattttagtatactagaaagtaataca 1370
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1307 CAGCCAGTCACATTTTTAAACAGTTACACTCTTATTTTAGTGTATACATAGAAAGTAATACA 1366

Qy   1371 ataaacatgtcaaac 1385
          |||||||||||||||
Db   1367 ATAAACATGTCAAAC 1381
```

```
RESULT       9

LOCUS        AC087434        160714 bp        DNA        linear    HTG 05-JAN-2001
DEFINITION   Pan troglodytes clone RP43-90F9, WORKING DRAFT SEQUENCE, 12
             unordered pieces.
ACCESSION    AC087434
VERSION      AC087434.1  GI:12039251
KEYWORDS     HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE       chimpanzee.
ORGANISM     Pan troglodytes
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE    1  (bases 1 to 160714)
AUTHORS      Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W.,
             Bouffard,G.G., Brinkley,C., Brooks,S., Dietrich,N.L., Granite,S.,
             Guan,X., Gupta,J., Ho,S.-L., Idol,J.R., Karlins,E., Lee-Lin,S.-Q.,
             Legaspi,R., Lin,M., Maduro,Q.L., Maduro,V.B., Masiello,C.,
             Mastrian,S.D., McCloskey,J.C., McDowell,J., Pearson,R., Prasad,A.,
             Shevchenko,Y., Snyder,B., Stantripop,S., Thomas,J.W., Thomas,P.J.,
             Tiongson,E.E., Touchman,J.W., Tsurgeon,C., Vogt,J.L., Walker,M.A.,
             Wetherby,K.D., Zhang,L.-H. and Green,E.D.
TITLE        NISC Comparative Sequencing Initiative
JOURNAL      Unpublished
REFERENCE    2  (bases 1 to 160714)
AUTHORS      Green,E.D.
TITLE        Direct Submission
JOURNAL      Submitted (05-JAN-2001) NIH Intramural Sequencing Center, 8717
             Grovemont Circle, Gaithersburg, MD 20877, USA
COMMENT      -------------- Genome Center

             Center: NIH Intramural Sequencing Center
             Center code: NISC
             Web site: http://www.nisc.nih.gov
             Contact: nisc_mouse@nhgri.nih.gov
             -------------- Project Information
             Center project name: aod
             Center clone name: 090F09
             -------------- Summary Statistics
             Sequencing vector: plasmid; n/a; 100% of reads
             Chemistry: Dye-terminator Big Dye; version 0.990319
             Assembly program: Phrap; version 0.990319
             Consensus quality: 152563 bases at least Q40
             Consensus quality: 154387 bases at least Q30
             Consensus quality: 155752 bases at least Q20
             Insert size: 137000; agarose-fp
             Insert size: 137000; pulse-field-gel
             Insert size: 155614; sum-of-contigs
             Quality coverage: 9.32x in Q20 bases; agarose-fp
             Quality coverage: 9.32x in Q20 bases; pulse-field-gel
             Quality coverage: 8.00x in Q20 bases; sum-of-contigs
             --------------
*  NOTE: This is a 'working draft' sequence. It currently
*  consists of 12 contigs. The true order of the pieces
*  is not known and their order in this sequence record is
```

```
*  arbitrary. Gaps between the contigs are represented as
*  runs of N, but the exact sizes of the gaps are unknown.
*  This record will be updated with the finished sequence
*  as soon as it is available and the accession number will
*  be preserved.
                1                2808:  contig of 2808 bp in length
             2809                2908:  gap of unknown length
             2909                6190:  contig of 3282 bp in length
             6191                6290:  gap of unknown length
             6291               12538:  contig of 6248 bp in length
            12539               12638:  gap of unknown length
            12639               20815:  contig of 8177 bp in length
            20816               20915:  gap of unknown length
            20916               28981:  contig of 8066 bp in length
            28982               29081:  gap of unknown length
            29082               40699:  contig of 11618 bp in length
            40700               40799:  gap of unknown length
            40800               49520:  contig of 8721 bp in length
            49521               49620:  gap of unknown length
            49621               62547:  contig of 12927 bp in length
            62548               62647:  gap of unknown length
            62648               76465:  contig of 13818 bp in length
            76466               76565:  gap of unknown length
            76566               98277:  contig of 21712 bp in length
            98278               98377:  gap of unknown length
            98378              127174:  contig of 28797 bp in length
           127175              127274:  gap of unknown length
           127275              160714:  contig of 33440 bp in length.
```

```
FEATURES             Location/Qualifiers
     source          1..160714
                     /organism="Pan troglodytes"
                     /db_xref="taxon:9598"
                     /clone="RP43-90F9"
                     /clone_lib="RP43"
     misc_feature    1..2808
                     /note="assembly_fragment"
     misc_feature    2909..6190
                     /note="assembly_fragment"
     misc_feature    6291..12538
                     /note="assembly_fragment"
     misc_feature    12639..20815
                     /note="assembly_fragment"
     misc_feature    20916..28981
                     /note="assembly_fragment"
     misc_feature    29082..40699
                     /note="assembly_fragment"
     misc_feature    40800..49520
                     /note="assembly_fragment"
     misc_feature    49621..62547
                     /note="assembly_fragment"
     misc_feature    62648..76465
                     /note="assembly_fragment"
     misc_feature    76566..98277
                     /note="assembly_fragment"
     misc_feature    98378..127174
                     /note="assembly_fragment"
     misc_feature    127275..160714
                     /note="assembly_fragment"
                     /clone_end:SP6
                     vector_side:right"
BASE COUNT    37810 a    40793 c    40289 g    40706 t    1116 others
ORIGIN
```

```
Qy      4 tttttttggagatggagtttcgctccttgttgcccaggctggagtgcaatggcgcaatct 63
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 132561 TTTTGTTTGAGATGGAGTTTTT-GCTCTTGTTGCCCAGGCTGGAGTGCAAGTGTAATGCACGATCT 132619

Qy     64 cagctcaccgcaacctccgcctccgggttcaagcgattcctcctgcctcagcctcccagg 123
```

Best Local Similarity 41.1%;   Score 592.2; DB 2;   Length 160714;
Query Match   71.2%;   Pred. No.1.6e-165;
Matches 944; Conservative   0; Mismatches   348; Indels   33; Gaps   11;

```
Db 132620  CAGCTCACCAACAACTTCCACCTCCCAGGTTCAAGCAAATTCTCCTACCTCAGCCTCCCAAG  132679
           |||||||| |||  |||||||||  ||||||||||   || |||||||||| ||||||||||
Qy    124  tagctgggattacaggcatgtgccaccacgctcggcctaatttgtattttttagtag  183
           ||||||||||||||||||||| |||||| ||  || |||||||| ||||||||    |||
Db 132680  TAGCTGGGATTACAGGCATGTGTCACCACACCGGCTAATTTTTGTATTTTT---AGT  132735

Qy    184  agatgggagttctccatgtgttggtcaggctgtgtctccgaacctccagatgatccct  243
           || ||||| ||||||||||| ||  |||| ||| ||  || |||||| ||| ||||| 
Db 132736  AGACGGGGTTTCTCCATGTTGTTCCCANGTGTGGTTCAGTTGGTTGGAACCCCGACCTCAGANGATCCGC  132795

Qy    244  ccgtcctcggcctcccaaaagtgctagatacaggacggccacgctctgcctgg  303
           || ||| ||||||||||||| ||||| ||||| || |||||| ||| |||||
Db 132796  CCACCTTGGCCTCCCAAAGTGCT-GGGATTACAGGTGTGAGTTTACANGAAAACATTTA  132854

Qy    304  ctaattttgtggttagaaaacagggtttcactgatgtgcccaagctggtctcctgagctca  363
           |     |     ||||  ||| || ||  |  |||  |||||||||    ||| ||  |||
Db 132855  AGTATGAAAGCAGTACCTCCGGAATATGCGGCGCCCTCCAAACCAGCGTTTTGAAGACTC  132914

Qy    364  agcagtccaacctgcctcagcctcccaaagtgctgggattacaggcgcgtgcagccgtg----  419
           || |||   |  |||| | |   |||   ||||||| | |  || |||  | | |||||
Db 132915  TTCACTTGTTCTGGTTGGGCCACCACCACTCAGCTCCCACTAAAAACTGGGCGTCGCCGCTCTTT  132974

Qy    420  -----------cctggccttttattttttattttttaagacaacaggtgtcccaactctta  468
                      || ||||| ||||||||||||||||| |  ||  || |||||||||| ||
Db 132975  CCCCCCCCATCCTCGCCTTTTTTTTTTTTTTTTTTTTGTT-GTCTTGTTTTGTTG  133033

Qy    469  cccaggagtgaagtgcagtggtgtgatcacagctcaagctcaactcctggagatc-  527
           ||||||  | ||||||||||||| |||| ||| || || || ||||||||||||
Db 133034  CCCAGGGCTGGAGTGCAGTGGTGCGAAACATGGCTCACTGCAGCCTTGATCTCTTGGGCTCA  133093

Qy    528  aagcatcctccctgcctcagcctcccaaagtagctgggaccaaaagaacatgcaaccactacaac  587
           ||| | ||||| | ||||||||||  |  |||||||| ||  |||| ||  ||| || ||||||
Db 133094  AAGAATCCTCCCACCTCAGCCTCCTGAGTAGCTGGGACTATAGGTGCGCCACCATGCC  133153

Qy    588  tggctaatttttattttttgtgagacagagtcctcaacttctgtcaccag  647
           | | ||||||||||||||| | ||| |||| ||  ||| ||| |||||||
Db 133154  CAACTAATTTTTTTTTTTTTTTTTTT--TTTTTGAGATGGGAGTTTCGCGTTCGTTGCCCAG  133211

Qy    648  gctcggagttcgcagttggcgcaatcttggcttcactgcaacctctgcctcccgggttcaagtta  707
           ||||| |  | |||  ||| || |||  | ||||||||||||||||||||||||||||||||  ||
Db 133212  GCTCGGGGTGCAATGGTGCAGTCTCGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGCAA  133271

Qy    708  ttctcctgcccccagcctcctgagtagctgggactacaggcgcccaccaccagcctagctaat  767
           |||||||||| ||| ||| | | |||||||||| ||||| |||| || ||| ||| |||| 
Db 133272  TTCTCCTGCCTCAGCCCCCCAAATAGCTGGGATTTCAGGCANCGGCATGGCCANCACCTGGCT  133331

Qy    768  ttttttgtattttttaagtagagatgggg--ttcaccatgttcgccaggttgatcttgatctc  826
           ||||| |||| | || | |||||||||||  ||||| |||| ||||||| |||||| ||||||
Db 133332  AATTTTTGTATTGTTATTGAGAGATGGGGTTTCACTATGTCAGTCAGGCTGGTCTCAAACTC  133391

Qy    827  tggacct--tgtgatctgcctgctcggcctcccaaacttggggattacaggcgtgagc  884
           |||| ||  ||||||  ||||||| ||||||||||| | || ||||||||||||||||
Db 133392  CTGACCTCAGAGAATCCTCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGC  133451

Qy    885  caccacgcccgg--cttaatttttaatttgttttgttgttaaatggaatccactctgttac  943
           || | |||| ||  || |||||||| |||| ||||||||| |  |||| || || |||| 
Db 133452  CACTGTGCCTGGCCTGCCCAATTTTTTTTTTGTTTTTTTTTGGAGACAAAGTCTCGCTCTGTTGC  133511

Qy    944  ccaggctggagtgcaatggcacaatctcggctcactgcaacctctccgggctcaa  1003
           ||||||||||||||||| ||||| ||||||||||| |||||||||| |||| |||
Db 133512  CCAGGCTGGAGTGCAAATGGCACCCTGCTCACTCTGCCTGCCTCCTCTCGGGTTCAA  133571

Qy   1004  gcgattactcctgtctcagcctcccaagcactgcacccacacc  1063
           ||||| | ||||| ||||||| ||||| | ||| |||| ||||
Db 133572  GCGATTCTCCTGCCTCAGCGTCCCAAGTAGCTGGGACTACAGGCGCATGCCACCACGCCC  133631

Qy   1064  cgctaattttgtattttcattagagacgggggtttcaccatgttcacctca  1123
           |||||||||||||||| | |||||| |   |||||||||||||  ||| ||
Db 133632  GGCTAATTTTTGTATTTTTAGTAGAGACGAGTTTCACCATGTTGGCCAGGCTGGTCTTG  133691

Qy   1124  aactcctgacctcaggtgacccacctcccaaagtgctgggattacaggcg  1183
           ||||||| ||||||| |   |||||||||||||||||||||||||||| |
Db 133692  AACTCCTAACTTCAGGTGGATCCACCCGGCTGCGTCCCAAAGTGCTGGGATTACAGGCA  133751

Qy   1184  tgagccacctcaccaccagcc-----ggctaattlagataaaaaatatgtagcaatgggggg  1239
           |||||||||| ||| ||||||     ||| |||| |||| |||||| |||||| ||||| 
Db 133752  TGAGCCACCACACCTGGCCTCTGCCCAACTAATTTTAAAAAATTTTGTAGAGACAGGA  133811

Qy   1240  tcttgctatgttgcccaggctggtctccaaacttctggcttcatgcaatcctccaaatga  1299
           || | ||||||||  |||||||||||| | | ||||| ||| | |||| |||||||| ||
Db 133812  TCTCACTATGTTGTCCAGACTGGTCTTGAACCTCCTGGGCTTAAGGGATCGTCCTGTCTCA  133871

Qy   1300  gccac  1304
           ||| |
Db 133872  GCCTC  133876
```

RESULT 10

```
LOCUS       AC005057               99370 bp    DNA     linear   PRI 30-SEP-2000
DEFINITION  Homo sapiens BAC clone CTB-52H6 from 7q11.2-q22, complete sequence.
ACCESSION   AC005057
VERSION     AC005057.2  GI:6587915
KEYWORDS    HTG.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 99370)
  AUTHORS   Sulston,J.E. and Waterston,R.
  TITLE     Toward a complete human genome sequence
  JOURNAL   Genome Res. 8 (11), 1097-1108 (1998)
  MEDLINE   9063792
REFERENCE   2  (bases 1 to 99370)
  AUTHORS   Dauphin,S., Stoneking,T. and Ahrens,C.
  TITLE     The sequence of Homo sapiens BAC clone CTB-52H6
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 99370)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JUN-1998) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   4  (bases 1 to 99370)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-DEC-1999) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   5  (bases 1 to 99370)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-SEP-2000) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
            On Dec 16, 1999 this sequence version replaced gi:3212934.
COMMENT     ----------------------  Genome Center
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site: http://genome.wustl.edu/gsc
            Contact: sapiens@watson.wustl.edu
            ----------------------  Summary Statistics
            Center project name: H_RG052H06
            ----------------------
            NOTICE:  This sequence may not represent the entire insert of this
            clone.  It may be shorter because we only sequence overlapping
            clone sections once, or longer because we provide a small overlap
            between neighboring data submissions.

            This sequence was finished as follows unless otherwise noted:
            all regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by sequence
            from more than one subclone; and the assembly was confirmed by
```

restriction digest.

MAPPING INFORMATION:
The sequence of this clone was established as part of a mapping and
sequencing collaboration between the NHGRI Chromosome 7 Mapping
Project (Eric D. Green, Director), John D. McPherson in the
Department of Genetics (Washington University), and the Washington
University Genome Sequencing Center. For additional information
about the map position of this sequence, see
http://www.nhgri.nih.gov/DIR/GTB/CHR7, send
mailto:egreen@nhgri.nih.gov , or see http://genome.wustl.edu/gsc

SOURCE INFORMATION:
Clone CTB-52H6 is from the first release of the human BAC library
CITB-978SK-B. The library contains cloned DNA from the male
fibroblast cell line 978SK. See: Shizuya et al., Proc. Natl.
Acad. Sci. USA 89:8794-7 (1992); U-J. Kim et al., Genomics 34:213-8
(1996). This clone is available from Research Genetics, Inc.
(http://www.resgen.com).
VECTOR: pBeloBAC11
Selection: chloramphenicol

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is CTA-270D13. Actual start of
this clone is at base position 135330 of CTA-270D13; actual end is
at base position 99370 of CTB-52H6.

FEATURES   Location/Qualifiers
source   1..99370
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7q11.2-q22"
                     /clone="CTB-52H6"
                     /clone_lib="CITB-978SK-B"
repeat_region   59..528
                     /rpt_family="Alu"
repeat_region   838..903
                     /rpt_family="Alu"
repeat_region   1428..1729
                     /rpt_family="MER2_type"
misc_feature   3190..3258
                     /rpt_family="Alu"
misc_feature   3193..3259
                     /note="match to EST AI208950 (NID:g3770892) gg60d12.x1"
misc_feature   3193..3259
                     /note="similar to Mus musculus
                     mc65g12.r1"
misc_feature   3193..3259
                     /note="match to EST AI208950 (NID:g3770892) gg60d12.x1"
misc_feature   3199..3224
                     /note="similar to Mus musculus
                     ms45b10.r1"
misc_feature   3587..3684
                     /note="match to EST AA333593 (NID:g1985887)"
misc_feature   3587..3684
                     /note="match to EST AL036193 (NID:g5927681)"
misc_feature   3588..3684
                     /note="match to EST AA076195 (NID:g1616065) zm18a06.r1"
misc_feature   3588..3684
                     /note="similar to Mus musculus EST W41925 (NID:g1325640)
                     mc65g12.r1"
misc_feature   3588..3684
                     /note="match to EST AI208950 (NID:g3770892) gg60d12.x1"
misc_feature   3588..3684
                     /note="similar to Mus musculus EST AA165787 (NID:g1744017)
                     ms45b10.r1"
misc_feature   3588..3684
                     /note="match to EST AA333593 (NID:g1985887)"
misc_feature   3769..3792
                     /note="match to EST AI208950 (NID:g3770892) gg60d12.x1"
misc_feature   3769..3833
                     /note="match to EST AL036193 (NID:g5927681)"
misc_feature   3769..3833
                     /note="match to EST AA333593 (NID:g1985887)"
misc_feature   3769..3833
                     /note="match to EST AL036193 (NID:g5927681)"

misc_feature   3769..3833
                     /note="similar to Mus musculus EST W41925 (NID:g1325640)
                     mc65g12.r1"
misc_feature   3769..3826
                     /note="match to EST AA076195 (NID:g1616065) zm18a06.r1"
misc_feature   3769..3833
                     /note="similar to Mus musculus EST AA165787 (NID:g1744017)
                     ms45b10.r1"
misc_feature   3770..3833
                     /note="match to EST AA852497 (NID:g2941090)"
misc_feature   3785..3833
                     /note="match to EST R58433 (NID:g828491)"
misc_feature   4726..4728
                     /note="match to EST AA076195 (NID:g1616065) zm18a06.r1"
misc_feature   5708..5895
                     /note="similar to Mus musculus EST W41925 (NID:g1325640)
                     mc65g12.r1"
misc_feature   5708..5895
                     /note="match to EST AA852497 (NID:g2941090)"
misc_feature   5708..5895
                     /note="match to EST AA333593 (NID:g1985887)"
misc_feature   5708..5879
                     /note="match to EST AL036193 (NID:g5927681)"
misc_feature   5716..5895
                     /note="match to EST R58433 (NID:g828491)"
misc_feature   5731..5896
                     /note="match to EST D59085 (NID:g968719)"
repeat_region   6070..6358
                     /note="similar to EST W28383 (NID:g1308331)"
repeat_region   6582..6695
                     /rpt_family="Alu"
repeat_region   6696..6717
                     /rpt_family="Alu"
repeat_region   6924..7013
                     /rpt_family="(A)n"
repeat_region   7014..7278
                     /rpt_family="MIR"
repeat_region   7279..7319
                     /rpt_family="Alu"
repeat_region   7637..7677
                     /rpt_family="MIR"
repeat_region   8214..8525
                     /rpt_family="L2_type"
repeat_region   8526..8811
                     /rpt_family="Alu"
repeat_region   8813..8866
                     /rpt_family="Alu"
repeat_region   9022..9059
                     /rpt_family="A-rich"
repeat_region   9729..9763
                     /rpt_family="(TATG)n"
repeat_region   10803..10908
                     /rpt_family="(CACAT)n"
repeat_region   11017..11137
                     /rpt_family="Alu"
repeat_region   11492..11762
                     /rpt_family="Alu"
repeat_region   11763..12059
                     /rpt_family="Alu"
repeat_region   12060..12093
                     /rpt_family="A-rich"
repeat_region   12331..12470
                     /rpt_family="L1"
repeat_region   12471..12597
                     /rpt_family="Alu"
repeat_region   12951..13298
                     /rpt_family="Alu"
repeat_region   13299..13572
                     /rpt_family="MER1_type"
repeat_region   13573..13857
                     /rpt_family="Alu"

Query Match   40.9%;   Score 590.4;   DB 9;   Length 99370;

Best Local Similarity 70.8%; Conservative 0; Pred. No. 5,3e-165;
Matches 94%; Indels 43; Gaps 10;

```
Qy      4 tttttttgagatggagtttcgctcttgttgcccaggctggagtgcaatggcgcaatct 63
            ||||| |||||||||||||| |||||||||| |||||||| ||||| |||| ||| ||
Db  69541 TTTTGTTTGAGATGGAGTTTT-GCTCTTGTTGTTGCCAGGCTGAGTGTAATGGCACGATCT 69599

Qy     64 cagctcaaccgcaacctccgcctcccgggttcaagcgattctcctgcctcagcctcccag 123
            ||||||| || ||||| ||| ||||| ||||||||| ||| |||||||| ||||||||||
Db  69600 CAGCTCACCACAACTTCGGCTCCCCAGGTTCAAGCAATTTCCTACCTCACCAGCCTCCCAG 69659

Qy    124 tagctgggattacaggc... 183
            |||||||||||||||
Db  69660 TAGCTGGGATTACAGGCANTGTCACCACCACGCCCAATTTTTGTANTTTT----AGT 69715

Qy    184 agatggagtttctccatgttggtcaggctggtctcgaactcccgacctcaggtgatcct 243
            |||| |||||||||||||||||||||| |||||| |||| |||| ||||||| |||||
Db  69716 AGACGGGGTTTCTCCATGTTGGTCAGGTTGGTCTTGAACCCCTGACCTCAGATGATCCGC 69775

Qy    244 ccgtctcggcctcccaaagtgctagatacagggcacgagcctgcctg 303
            ||| |||||||||||||||||||| |  |||||| ||| ||| |||
Db  69776 CCACCTTGGGCCTCCCAAAGTGCT-GGGATTACAGGCGTGAGTTTTACAAGAAAACATTTA 69834

Qy    304 ctaattttttggttagaaaacagggttcactgatgtgtccccaagctggtctcctgagctca 363
            |||||||   ||  |||||  ||| ||    ||||||    ||| ||| ||||| ||| ||
Db  69835 AGTATGAAAGCAGTACCTCCGGAANTANTGCACGGCCCTTCAAAACCAGCCTCTGAGACTC 69894

Qy    364 agcagtccaactgcttcagcctcccaaagtgctggattacaggcgtgaagccgtgcctg 423
            |  || ||  |  || ||| ||||||||  |||  || || |   ||  | |||| |||
Db  69895 TTCACTTGTCCTCGGCCCACCACTCAAAACTGGGCGTGTGGTTTGTCTT 69954

Qy    424 gcc------...-------ttttattttttattttttaagacacaggtgtcccactc 465
            |||                ||||| |||||||||||| | |||| || |||||| ||
Db  69955 CCCCCCTCCCCATCCCTGCCTTTTTTTTTTTTTTTTTTTTTTTTGAGACAGAGTCTCGTTTG 70014

Qy    466 ttaccccaggatgaagtgcagtggtgtgatccacagctcactgcagccttcaactcctgaga 525
            ||  |||||| | || |||||||| |  ||||| |||||||||| ||||  ||| |||||
Db  70015 TTGCCCAGGCTGGAGTGCAGTGGTGCAAACATGGCTCACTGCAGCCTCGAGCTCCTGGGC 70074

Qy    526 tc-aagcatcctcctgcctcagcctcccaagtagctgggaccaaagacatgcaccactac 584
            || ||| | |||| |||||||||||||||||||||||||| | || || ||| |||||||
Db  70075 TCAAAGGATCCTCCCACTCAGCCTCTGTGAGTGAGCTGGGACTATAGGTGCATGCCACCAT 70134

Qy    585 acctggctaattttttattttttattttttgagacagagtctcaacttctgtcacc 644
            | ||  |||||||||| ||||||   |||||||| ||||||||  | |||||| |||
Db  70135 GCCCAACTAATTTTTTTTTTTTT-------TTTTTTGAGATGGAGTTTCGCTCTGTTGTTGCC 70186

Qy    645 caggctggagtgcgagtggcgcaatcttggctcactgcaacctctgcctcccgggttcaag 704
            |||||||| |||||| |||||||||||| ||||||||||||||| ||||| |||||||||||
Db  70187 CAGGCTGGGGTGCGAAATGGCGCAATCTCGGCTCACTGCAACCTCCGCCTCCGGGTTCAAG 70246

Qy    705 ttattcctgcccagcctcctgagtagctgggactacaggcgcccaacccagcctagct 764
            |  || ||||| |||||||| | ||||||||| |||||||| |  |  |||||| |||
Db  70247 CAATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGCATGCGCCACCACGCCTG 70306

Qy    765 aattttttgtattttagtagagacgatgggg-ttcaccatgttgtccaggttgatcttgat 823
            |||||||| ||| || ||||||| ||| || ||| |  ||||| ||||| ||
Db  70307 GCTAATTTTTGTATTTGTAGAGAGGTTTCAGTCAGTCCTGTTGTCTCAAA 70366

Qy    824 ctctggacct--tgtgatctggctcggctcggcctcccaaagtgctggattacaggcgtg 881
            |||| |||||  ||| || ||||||||||||||| |||||||||||| || |||| ||||
Db  70367 CTCCTGACCTCAGGAGATCCTCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTG 70426

Qy    882 agcaccaacgcccggctt--attttttaattttgttgtttgaaatggaatctcactct 938
            |||||  | ||||| ||| ||||||||||||  |||||||| || || |||||| |||
Db  70427 AGCCACTGTGCCCGGCCTCCCAAATTTTTTTTTTTATTTTTGTTGTTTGGGAGACAGAGTCTCGCT 70486

Qy    939 gttaccagggtggagtgcaaatggcgcaaatctcgcctcactgcaacctctgcctcccggg 998
            |||||| ||| |||||||  ||||| ||||| |||||||||||||||||| ||||||||||
Db  70487 CTTTCCCAGGCTGGAGTGGGAAATGGGCAACCACCTCTGCCTCCCGG 70546

Qy    999 ctcaaggcgattcctcgtctcccaagcaagcagctgggattacgggcacctgcgcacca 1058
            |||||| | ||||| |||||||||||  ||||||||||||||| ||||||||| |||||
Db  70547 TTCAAGCGGATTCTCCTGCCTCAGTCTCCCAAGTAGCTAGGACTACAGGCATGTGCCACCA 70606

Qy   1059 cacccccgctaattttttgtatttttcattagagacgggggtttcaccatattctcaggctgg 1118
            ||||| | |||||||||||||| ||| |||||||| |||||||||||||| || |||| |||
Db  70607 CGCCCGGCTAATTTTTTGTATTTTTAGTAGAGACGAGGTTTCACCATGTTGGCCAGGACTGG 70666

Qy   1119 tctcaaactcctgaccttgaggtgacccacctgccttggcctcccaaagtgctggattac 1178
            ||| |||||||| ||||| ||||||| |||| ||| ||||||||||||||||||| |||||
Db  70667 TCTTGAACTCCTAACTTCAGGTGATCACCCGCCTCGGCCTCCCAAAGTGCTGGGATTAC 70726

Qy   1179 aggcgtgagccacctcacccagcc-----ggctaatttagataaaaaatatatgtagcaatg 1234
            ||||  ||| |||| || |||||| |     ||| |||  || ||||||| |  ||||| ||
Db  70727 AGGCATGAGCCACCACACCCAGCCGCCTCTCCCAACTAATTAAAAAAAATTTTTGTAGAGA 70786

Qy   1235 gggggtcttgctatgttgcccaggctggtctcaaactctgggggctcaatcctcca 1294
            ||||| ||||| || ||||||||| ||||||| ||| |||||| ||||  | ||||
Db  70787 CAGGANTCTCACTANTGTTGTCCAGACTGGTCTTGAACTCCTGGGGCTCAAGGGATCGTCTG 70846

Qy   1295 aatgagccac 1304
            ||||||||||
Db  70847 TCTCAGCCTC 70856
```

RESULT 11
LOCUS   AX330738/c   65608 bp   DNA   linear   PAT 09-JAN-2002
DEFINITION Sequence 1247 from Patent WO0194629.
ACCESSION   AX330738
VERSION   AX330738.1 GI:18103716
KEYWORDS   .
SOURCE   human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (sites)
AUTHORS   Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
            Horrigan,S., Soppet,D.R. and Weaver,Z.
TITLE   Cancer gene determination and therapeutic screening using signature
            gene sets
JOURNAL   Patent: WO 0194629-A 1247 13-DEC-2001;
            Avalon Pharmaceuticals (US)
FEATURES   Location/Qualifiers
     source   1..65608
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT   15987 a 16979 c 16761 g 15878 t   3 others
ORIGIN

Query Match 40.8%; Score 587.8; DB 6; Length 65608;
Best Local Similarity 70.8%; Conservative 0; Pred. No. 3e-164;
Matches 940; Mismatches 347; Indels 44; Gaps 10;

```
Qy      4 tttttttgagatggagtttcgctcttgttgcccaggctggagtgcaatggcgcaatct 63
            ||||| |||||||||||||| |||||||||| |||||||| ||||| |||| ||| ||
Db  39993 TTTTGTTTGAGATGGAGTTTT-GCTCTTGTTGTTGCCAGGCTGAGTGTAATGGCACGATCT 39935

Qy     64 cagctcaaccgcaacctccgcctcccgggttcaagcgattctcctgcctcagcctcccag 123
            ||||||| || ||||| ||| ||||| ||||||||| ||| |||||||| ||||||||||
Db  39934 CAGCTCACCACAACTTCGGCTCCCCAGGTTCAAGCAATTTCCTACCTCACCAGCCTCCCAG 39875

Qy    124 tagctgggattacaggcgtgcaccaccacgctcggctaattttt 183
            |||||||||||||||   |||||||||||| |||  |||||||||
Db  39874 TAGCTGGGATTACAGGCATGTCACCACCACGCCCAATTTTTGTANTTTT----AGT 39819

Qy    184 agatggagtttctccatgttggtcaggctggtctcgaactcccgacctcaggtgatcct 243
            |||| |||||||||||||||||||||| |||||| |||| |||| ||||||| |||||
Db  39818 AGACGGGGTTTCTCCATGTTGGTCAGGTTGGTCTTGAACCCCTGACCTCAGATGATCCGC 39759

Qy    244 ccgtctcggcctcccaaagtgctagatacagggcacgagcctgcctg 303
            ||| |||||||||||||||||||| |  |||||| ||| ||| |||
Db  39758 CCACCTTGGGCCTCCCAAAGTGCT-GGGATTACAGGCGTGAGTTTTACAAGAAAACATTTA 39700
```

```
RESULT 12
LOCUS        AX332242/c
DEFINITION   AX332242
             Sequence 2751 from Patent WO0194629.
ACCESSION    AX332242
VERSION      AX332242.1  GI:18122876
KEYWORDS     .
SOURCE       human.
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (sites)
  AUTHORS    Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
             Horrigan,S., Soppet,D.R. and Weaver,Z.
  TITLE      Cancer gene determination and therapeutic screening using signature
             gene sets
  JOURNAL    Patent: WO 0194629-A 2751 13-DEC-2001;
             Avalon Pharmaceuticals (US)
FEATURES             Location/Qualifiers
     source          1..55608
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT       15987 a   16979 c   16761 g   15878 t      3 others
ORIGIN
```

```
Query Match   40.8%;  Score 587.8;  DB 6;  Length 65608;
Best Local Similarity   70.6%;  Pred. No. 3e-164;
Matches  940;  Conservative    0;  Mismatches  347;  Indels   44;  Gaps  10;
```

Qy   304 ctaattttgtggttagaaacagggttcactgatgtgcccaagctggtctcctgagctca 363
Db 39699 AGTATGAAAAGCAGTACCTCCGGAATAATGCACGGCCCTTCCAAACCCAGCCTCTGAAGACTC 39640

Qy   364 agcagtccaactgcctcagcctccgaagtgctgggatta------------------- 403
Db 39639 TTCACTTGTCTCTGGCTGTGGGCCCTCACCTCACCTAAAACTGGGCTGTGTTCTCTT ------ 39580

Qy   404 caggcgtgcagccgtgcctgggccttttatttttttatttttaagacacaggtgtcccac 463
Db 39579 CCCCCCTCCAtCCCCTGCCCTTTTTTATTTTTTTATTTTTTGAGACAGAGTCTTGTTT 39520

Qy   464 tcttacccaggatgaagtgcagtgtgtgatcacagcctcaactcctcctga 523
Db 39519 TGTTGCCCAGGCTGGAGTGCAGTGGCTCAAACAtGCCACCTTGATCTCTTGG 39460

Qy   524 gatc-aagcatcctcctgcctcagcctcccaaagtagctgggaccaaagacatgcaccact 582
Db 39459 GCTCAAAAGGATCCTCCACCTCAGCCTCTGAGTAGCTGGGACTATAGGTGCCATGCCACC 39400

Qy   583 acacctggctaattttt tattttttaagacagagagtctcaactctgtca 642
Db 39399 ATGCCCAACTAATTTTTTTTTTTTTTT------- TTTTTTGAGATGGAGTTTCGCTCTTGTTG 39348

Qy   643 cccaggctggagtgcagtggcgcaaatcttggctcactgcaacctctgcctcccgggttca 702
Db 39347 CCCAGGCTGGGGTGCAATGGTGCGAATCTCGGCTCACTGCAACCTTCCTGGGTTCA 39288

Qy   703 agttattcctcctgcccccaggcctcctgagttagctggactacaggcgtgcgccaccacgcctag 762
Db 39287 AGCAAATTCTCCTGCCTCAACCCCCCAAGTAGCTGGGATTTACAGGCATGCCACCACC 39228

Qy   763 ctaattttttgtatttttagtagagatgggg-ttacaccatgttcgtccaggttgatcttg 821
Db 39227 TGGCTAATTTTGTATTGTAGTAGAGATGGGGTTTCACTATGTCAGTGGCCATCACC 39168

Qy   822 atcctctggggacct--tgtgatctgcctcggcctcccaaagtgctgggattacaggcg 879
Db 39167 AACTCCTGACCTCAGGAGATCCTCCTGCCTCGGCCTCCCAGAGGGCTGGGATTACAGGTG 39108

Qy   880 tgagccacaccagcccgcgcctt--attttttaattttgttgttgaaatgaatcctcactc 937
Db 39107 TGAGCCACTGTCCTGCCTGGCCCAATTTTTGTTTTTTTTTTGGAGACAAAGTCTCGC 39048

Qy   938 tgttacccaggctggagtgcaaatggccaaatctcgtgcctccccgg 997
Db 39047 TCTTTCCCAGGCTGGAGTGGAAATGGCACCCTCGGCGTCCCCGAGATGAtCCGC 38988

Qy   998 gctccaagcgattctcctgtcttcagcctcccaagacgagctgggattacgggcaacctgccacc 1057
Db 38987 GTTCAAGCGATTCTCCAGCCTCCCAAAGTGCT-GGGATTACAGGCGTGGAGCCACCACATTTA 38928

Qy   1058 acacccgctaattttttgtattttcatttagaaggcggggtttcaccatatttgtcaggctg 1117
Db 38927 ACGCCCGGCTAAtTTTTTGTAttTTTAGtAGAGACGAGTTTCACCAtGtTGGCCAGACTG 38868

Qy   1118 gtctcaaactcctgacctcaggtgacccaccactgcctcagcgcttccaaagtgctgggatta 1177
Db 38867 GTtCTGAACTCCTAACTTCAGGtGAtCACCACGCCCTGCTCCAGCGcctcaaaagTGGGATTA 38808

Qy   1178 caggcgtgagccacacctcacccagcc---ggctaatttagatataaaatatgtgagcaat 1233
Db 38807 CAGGCAtGAGCCACCACTGGCCTCCCAACTAAtTAAAAAAAAATTTTTTGTAGAG 38748

Qy   1234 gggggttcttgctatgttgcccaggctgtctcaaaactctgtgcttcatcgcaatcctcc 1293
Db 38747 ACAGGAGTCACATATGTTGTCCAGACTGGTCTCTTGAAACTCCTGGGCTTGAAGGAtCGTCCT 38688

Qy   1294 aaatgagccac 1304
Db 38687 GTCTCAGCCTC 38677

```
RESULT 13
LOCUS       AX335496              65608 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 6005 from Patent WO0194629.
ACCESSION   AX335496
VERSION     AX335496.1  GI:18126215
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
            Horrigan,S., Soppet,D.R. and Weaver,Z.
  TITLE     Cancer gene determination and therapeutic screening using signature
            gene sets
  JOURNAL   Patent: WO 0194629-A 6005 13-DEC-2001;
            Avalon Pharmaceuticals (US)
FEATURES             Location/Qualifiers
     source          1..65608
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT    15987 a  16979 c  16761 g  15878 t      3 others
```

ORIGIN

```
Query Match   40.8%;  Score 587.8;  DB 6;  Length 65608;
Best Local Similarity  70.6%;  Pred. No. 3e-164;
Matches  940;  Conservative  0;  Mismatches 347;  Indels  44;  Gaps  10;
```

Qy   643 cccaggctggagtgcagtggcgccaatcttggctcactgcaacctctgcctcccgggttca 702
Db   39347 CCCAGGCTGGGGTGCAATGGTGCGGATCTCGGCTCACTGCAACCTTCTGCCTTCCGGGTTCA 39288

Qy   703 agttattctctgcccccagcctcctgagtagctgggactacaggcgcccaccacgccctag 762
Db   39287 AGCAAATTCTCCTGCCTCAGCCTCCCAAACCCCCAGTAGCTGGGATTACAGGCGTGCCATCACACC 39228

Qy   763 ctaatttttgtgtattttttagtagagatgggg--ttcaccatgttcgccaggttgatcttg 821
Db   39227 TGGGCTAATTTTGTATTGTTAGTAGAGAIGGGGGTTTCACTATGTCAGGCGCCGTTCTCA 39168

Qy   822 atctctggacct--tgtgatctgcctcgcctcccaaagtgctgggattacaggcg 879
Db   39167 AACTCCTGACCTCAGGAGGATCCTCCTGCCTCGGCCTCCCAGAGTGCTGGGATTACAGGCG 39108

Qy   880 tgagccacacgcccggctt--attttttaatttgtttgttgaatctcactc 937
Db   39107 TGAGCCACTGTGCCTGGCCTGCCAANTTTTTTTTTTTTTTTTTGGGACAAAGTCTCGC 39048

Qy   938 tgttacccaggctggagtgcaatggccaaatctcggctcaactgcaacctctgcctcccgg 997
Db   39047 TCTTTCCCAGGCTGGAGTGGAAATGGCCACCCTCGGCTCAACCTTCTGCCTTTCGG 38988

Qy   998 gctcaagcgattctcctgtctcagcctcccaaagtagctgggacctgccacc 1057
Db   38987 GTTCAAGCGGATTCTTCCTGCCTCCAGTCTCCCAAGTAGCTACAGGCCATGTGCCACC 38928

Qy   1058 acaccccgctcaatttttgtattttcacacatattgtcaaggctg 1117
Db   38927 ACGCCCGGGTCTAATTTTTGTGTATTTTTTAGTAGAGACAAGGTTTCACCATGTTGGCCAGGACTG 38868

Qy   1118 gtctcaaactcctgacctcaaggtgaccaacctgcctcagcccttccaaagtgctgggatta 1177
Db   38867 GTCTTGAACTCCTACTTCAGGTGATCCACCCGCCTCCCAAAAGTGCTGGGATTA 38808

Qy   1178 caggcgtgagccacctcaccagcc---ggctaatttagataaaaaatatgtagcaat 1233
Db   38807 CAGGCAATGAGCACCACCACCTGGCCTGGCCTTGCCCAACTACAATTAAAAAAATTTTGTAGAG 38748

Qy   1234 gggggtcttgctatggtggccaggctggtctccaaacttctggcttcatgcaatcctcc 1293
Db   38747 ACAGGANTCTCACTATGTTGTCCAGACTGGTCTTGAACTCCTGGGCTTAAGGATCGTCCT 38688

Qy   1294 aaatgagccac 1304
Db   38687 GTCTCAGCCTC 38677
```

Qy        4  tttttttttgaagtggagtttcgctcttgttgcccaggctggagtgcaatggcgcaatct  63
Db    39993  ||||||||||||||||||||||||||||||||||||||||||| ||||||| |||||
           CAGCTCACCGCAACCTCCGCCTCCCGGGTTCCAAGCGATTCTCCTGCCTCAGCCTCCCAAG  39935

Qy       64  cagctcaccgcaacctccgcctcccgggttccaagcgattctcctgcctcagcctccc  123
Db    39934  CAGCTCACCACAACTTCCGGCCTCCAAGTAGCTGGGATTACAGGCGTGCCATCACACC  39875

Qy      124  tagctggattacaggcatgtgcaccacgccccaggttaatttttgtatttttttagtag  183
Db    39874  TAGCTGGGATTACAGGCAGTGTCACCACACCTGGGCTAATTTTTGTAITTTTT---AGT  39819

Qy      184  agatggagtttctccatgttggtcaggctggtctcgaactcctgacctcaggtgatcct  243
Db    39818  AGACGGGGTTTCTCCATGTTGGGCAGGGTCTCGAGTTTTTACAAGAAAACATTTA  39759

Qy      244  ccgtctcgggcctccccaagtgctagatacagggccaccaatgcccggctctgctgg  303
Db    39758  CCACCTTGGGCCTCCCAAAGTGCT-GGGATTACAGGTGTGAGTGTGAGTTTTACAAGAAAACATTTA  39700

Qy      304  ctaatttttgtggttagaaacagggtttcactgatgtgtgccaaggtggtctcctgagctca  363
Db    39699  AGTATGAAAAGCAGTACCTCCGGGAATATGCACAGCGCCCTCCAAACCCAGCGCCTCGAAGACTC  39640

Qy      364  agcagtccaactgcctcagccccaaagtgctgggatta------  403
Db    39639  TTCACTTGTCCTGGCGTGGGCGGCCACCACCTCCAGCTCCACTAAAAAACTGGGG  39580

Qy      404  caggcgtgcagccgtgcctgggccttttattttttattttttagcac  463
Db    39579  CCCCCCTCCANCCCTGCCTTTTTNTTTTTTTTTTGAGACAGAGTTTTG  39520

Qy      464  tcttacccaggatgaagtgcagtggtgtgatcacagctcactgcagcctcaactcctga  523
Db    39519  TGTTGCCCAGGCTGGAGTGCAGTGGCGCAAACANGGCTTCACTGCAACCTCTGCTTCCTGG  39460

Qy      524  gatc-aagcatcctcctgcctcagcctcccaagtagctgggaccaaagacatgcaccact  582
Db    39459  GCTCAAAGGATCCTCCACCCAGCCTCCGAGTAGCTGGGACTRTAGGTGCATGCCACC  39400

Qy      583  acacctggctaatttttattttttatttttagacaagtctcaaacctctgtca  642
Db    39399  ATGCCCAACTAANTTTTTTTTTTTT------TTTTTGAGAIGGAGTTTCGCTCTTGTTG  39348

Qy      643  cccaggctggagtgcaatggcgcaatcttggctcactgcaacctctgcctcccgggttca  702
Db    39347  CCCAGGCTGGGGTGCAATGGTGCGGATCTCGGCTCACTGCAACCTTCTGCCTTCCGGGTTCA  39288

Qy      703  agttattctctgcccccagcctcctgagtagctgggactacaggcgcccaccacgccctag  762
Db    39287  AGCAAATTCTCCTGCCTCAGCCTCCCAAACCCCCAGTAGCTGGGATTACAGGCGTGCCATCACACC  39228

Qy      763  ctaatttttgtgtattttttagtagagatgggg--ttcaccatgttcgccaggttgatcttg  821
Db    39227  TGGGCTAATTTTGTATTGTTAGTAGAGAIGGGGGTTTCACTATGTCAGGCGCCGTTCTCA  39168

Qy      822  atctctggacct--tgtgatctgcctcgcctcccaaagtgctgggattacaggcg  879
Db    39167  AACTCCTGACCTCAGGAGGATCCTCCTGCCTCGGCCTCCCAGAGTGCTGGGATTACAGGCG  39108

Qy      880  tgagccacacgcccggctt--attttttaatttgtttgttgaatctcactc  937
Db    39107  TGAGCCACTGTGCCTGGCCTGCCAANTTTTTTTTTTTTTTTTTGGGACAAAGTCTCGC  39048

Qy      938  tgttacccaggctggagtgcaatggccaaatctcggctcaactgcaacctctgcctcccgg  997

```
RESULT 14

LOCUS       HSU62293/c
HSU62293/c

LOCUS       HSU62293                 65608 bp    DNA     linear   PRI 26-OCT-1996
DEFINITION  Human LIM-kinase1 and alternatively spliced LIM-kinase1 (LIMK1)
            gene, complete cds.
ACCESSION   U62293
VERSION     U62293.1  GI:1432163
KEYWORDS    Williams syndrome.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 65608)
  AUTHORS   Frangiskakis,J.M., Ewart,A.K., Morris,C.A., Mervis,C.B.,
            Bertrand,J., Robinson,B.F., Klein,B.P., Ensing,G.J., Everett,L.A.,
            Green,E.D., Proschel,C., Gutowski,N.J., Noble,M., Atkinson,D.L.,
            Odelberg,S.J. and Keating,M.T.
  TITLE     LIM-kinase1 hemizygosity implicated in impaired visuospatial
            constructive cognition
  JOURNAL   Cell 86 (1), 59-69 (1996)
  MEDLINE   96291399
REFERENCE   2  (bases 1 to 65608)
  AUTHORS   Frangiskakis,J.M., Odelberg,S.J., Atkinson,D.L. and Keating,M.T.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-JUN-1996) Human Genetics, Univ. of Utah, 10 N 2030 E,
            Bldg 533, Suite 2100, Salt Lake City, UT 84112, USA
FEATURES             Location/Qualifiers
     source          1..65608
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7q11.23"
                     /tissue_type="placenta; hippocampus"
                     /note="Stratagene catalog No. 936205; chromosome
                     7-specific flow-sorted cosmid library from Lawrence
                     Livermore National Laboratories"
     gene            join(<1..101,1789..1885,12663..12801,13121..13230,
                     15072..15278,21920..22025,22122..22288,23055..23238,
                     23915..24001,24949..25080,27692..27751,27977..28042,
                     31846..32002,36610..36665,36919..37076,37166..38524)
                     /gene="LIMK1"
     mRNA            join(<1..101,1789..1885,12663..12801,13121..13230,
                     15072..15278,21920..22025,22122..22288,23055..23238,
                     23915..24001,24949..25080,27692..27751,27977..28042,
                     31846..32002,36610..36665,36919..37076,37166..38524)
                     /gene="LIMK1"
                     /product="LIM-kinase1"
                     /note="Williams syndrome region; Kiz-1; protein kinase
                     with two LIM domains"
     CDS             join(47..101,1789..1885,12663..12801,13121..13230,
                     15072..15278,21920..22025,22122..22288,23055..23238,
                     23915..24001,24949..25080,27692..27751,27977..28042,
                     31846..32002,36610..36665,36919..37076,37328)
                     /gene="LIMK1"
                     /note="Williams syndrome region; Kiz-1; protein kinase
                     with two LIM domains"
                     /codon_start=1
                     /product="LIM-kinase1"
                     /protein_id="AAB17545.1"
                     /db_xref="GI:1432164"
                     /translation="MRLTLLCCTWREERMGEEGSELPVCASCGQRIYDGQYLQALNAD
                     WHADCFRCCDCSASLSHQYYEKDGQLFCKKDYWARYGESCHGCSEQITKGLVMVAGEL
                     KYHPECFICLCGTFIGDGDGTYLVEHSKLYCGHCYQYQTVVTPVIEQILPDSPGSHLP
                     HTVTLVSIPASSHGKRGLSVSIDPPHGPPGCGTEHSHTVRQVGPGCMSPDVKNSIH
                     VGDRLLEINGTPIRNVPLDEIDLLIQETSRLLQLTLEHDPHDTLGHGLGPETSPLSSP
                     AYTPGSEAGSSARQRPVLRCSIDRSPGAGSLGSPASQRKDIGRSESLRVVCRPHIF
                     RPSDLIHGEVLGKGCFGQAIKVTHRETGEVMVMKELIRFDEETQRTFLKEVKVMRCLE
                     HPNVLKFIGVLYKDKRLNFITEYIKGGTLRGIIKSMDSQYPWSQRVSFAKDIASGMAY
                     LHSMNIIHRDLNSHNCLVRENKNVVADFGLARLMVDEKTQPEGLRSLKPDRKRYT
                     VVGNPYWMAPEMINGRSYDEKVDVFSFGIVLCEIIGRVNADPDYLPRTMDFGLNVRGF
                     LDRYCPPNCPSFYIPTIVRCCLDLDPEKRPSFVKLEHWLETLRMHLAGHLPLGPLQLEQL
                     DRGFWETYRRGEGSGLPAHPEVPD"

    CDS             join(89..101,1789..1885,12663..12801,13121..13230,
                     15072..15278,21920..22025,22122..22288,23055..23238,
                     23915..24001,24949..25080,27692..27751,27977..28042,
                     31846..32002,36610..36665,36919..37076,37328)
                     /gene="LIMK1"
                     /note="Williams syndrome region; Kiz-1; protein kinase
                     with two LIM domains; alternative translation initiation
                     site"
                     /codon_start=1
                     /product="alternatively spliced LIM-kinase1"
                     /protein_id="AAB17546.1"
                     /db_xref="GI:1432165"
                     /translation="MGEEGSELPVCASCGQRIYDGQYLQALNADWHADCFRCCDCSAS
                     LSHQYYEKDGQLFCKKDYWARYGESCHGCSEQITKGLVMVAGELKYHPECFICLTCGT
                     FIGDGDTYLVEHSKLYCGHCYQYQTVVTPVIEQILPDSPGSHLPHTVTLVSIPASSHG
                     KRGLSVSIDPPHGPPGCGTEHSHTVRQVGVDPGCMSPDVKNSIHVGDRLLEINGTPIR
                     NVPLDEIDLLIQETSRLLQLTLEHDPHDTLGHGLGPETSPLSSPAYTPGSEAGSSARQ
                     KPVLRCSIDRSPGAGSLGSPASQRKDIGRSESLRVVCRPHIFRPSDLIHGEVLGKGKG
                     CFGQAIKVTHRETGEVMVMKELIRFDEETQRTFLKEVKVMRCLEHPNVLKFIGVLYKD
                     KRLNFITEYIKGGTLRGIIKSMDSQYPWSQRVSFAKDIASGMAYLHSMNIIHRDLNSH
                     NCLVRENKNVVADFGLARLMVDEKTQPEGLRSLKPDRKRYTVVGNPYWMAPEMIN
                     GRSYDEKVDVFSFGIVLCEIIGRVNADPDYLPRTMDFGLNVRGFLDRYCPPNCPSFY
                     IPTIVRCCLDLDPEKRPSFVKLEHWLETLRMHLAGHLPLGPLQLEQLDRGFWETYRRGESG
                     LPAHPEVPD"

     polyA_site      38524..38529
                     /gene="LIMK1"
     polyA_site      38529..38534
                     /gene="LIMK1"
     misc_feature    6161
                     /gene="LIMK1"
                     /note="K2049 deletion breakpoint"
BASE COUNT    15987 a  16979 c  16761 g  15878 t      3 others
ORIGIN
```

Query Match   40.8%; Score 587.8; DB 9; Length 65608;
Best Local Similarity   70.6%; Pred. No. 3e-164;
Matches   940; Conservative   0; Mismatches 347; Indels 44; Gaps 10;

```
Qy          4 tttttttgagatgggagtttccgctcttgttgcccaaggcgtggggagtggcagcaatct 63
              |||| ||||||||||||||||| ||||||||||||||||||||||||| ||||||| |||
Db      39993 TTTTTTGTTTGAGATGGAGTTTT-GCTCTTGTTGCCCAGGCTGAAGTGTAATGGCAGATCT 39935

Qy         64 cagctcaccgcaacctccgcctccccggggttcaagcggattctcctgcctcccccag 123
              |||| |||||||||| | ||||| |||||| |||||||||||||||| ||||||||||||
Db      39934 CAGCTCACCACAACTTCCGCCTCCCAGGTTCAAGCGAACAATTCTCCTCTACCTCCCCCAG 39875

Qy        124 tagctgggattacaggcatgcacccacgctcggctaaatttgtatttttttagtag 183
              ||||||||||||||||||||||||||||||||| |||| | ||||||||||||| |||
Db      39874 TAGCTGGGATTACAGGCATGTGCCACCACACCTGGCTAATTTTTTGTATTTTT---AGT 39819
```

```
Db  38747 ACAGGATCTCACTATGTTGTCCAGACTGGTCTTGAACTCCCTGGGCTTAAAGGGGATCGTTCCT 38688
           || ||| ||||||||||||||| ||| ||||||||||||| ||| || | ||||| ||| |
Qy   1294 aaatggagccac 1304
           ||| |||
Db  38687 GTCTCAGCCTC 38677


Qy    184 agatggagtttctccatgttggtcaaggtggtctcgaactcccgacctcagatgatccct 243
           ||| || |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  39818 AGACGGGGTTTCTCCATGTTGGTCAAGGTGGTCTCGAACTCCCGACCTCAGATGATCCGC 39759

Qy    244 ccgtctcggcctcccaaagtgctagatacaagaggctgggttgcttcagatgatccctgg 303
           || || ||||||||||||||||  |||| |||||| |||||||| |||   |||  |
Db  39758 CCACCTTGGCCTCCCAAAGTGCT-GGGATTACAGATGTGAGTTTTACAAGAAAACATTTA 39700

Qy    304 ctaattttttgtgtagaaacagggtttcactgtgtccccaagatggtctctctgagctca 363
           |||   || || ||| ||| || ||| |||     ||    ||| || | | ||| ||
Db  39699 AGTATGAAAGCAGTACCTCCGGAATATGCACGGCCCTCCAAACCCAGCCCTCTGAAGACTC 39640

Qy    364 agcaggtcccacctgcctcagcctcccaaagtgctgggatta-------------------- 403
           | | | |||  |||||||||||||| ||||||||||||   |
Db  39639 TTCACTTGTCCTTGGCCTGGGCCGCCACCACCTCAGCCTCAAAACCTGGGCTGTGTTCTT 39580
                                                                       403

Qy    404 caggcggtgcaagccgtgcctggcccttttatttttttaagacacaggtgtccac 463
           |     ||| |   |||  ||| ||| ||||||||||||  |||| |||| | ||
Db  39579 CCCCCTCCATCCTGCCTTGTTTTTTTTTTTTTTTTTTTTGAGACAGAGTCTGTTT 39520

Qy    464 tcttacccaggatgaagtgcagtgtgtgatcacagctcagctcagctcaactcctga 523
           | || |||||| ||||||||||||| |||||||||| | |  || |||||||| |
Db  39519 TGTTGCCCAGGCTGGAGTGCAGTGGTCGAAACATGGCTCACTGCAGCCTCTTGG 39460

Qy    524 gatc-aagcatcctcctgcctcagcctcccaagtagctgggaccaaagacatgcaccact 582
           |||| | |||| ||| || ||||||||||| |||||||||  |  | | | ||| |||
Db  39459 GCTCAAAGGATCCTCCCACCTCAGCCTCCTGAGTAGCTGGGATTAGGTGCATGCCACC 39400

Qy    583 acacctggctaatttttattttttattttttgtagagacaggtttcaccatgttggtccac 642
           | ||| ||||||||||| || ||| |    | || || || ||||| | ||| |||  |
Db  39399 ATGCCAACTAATTTTTTGTTTTTT----------TTTTTGATGGGAGTTTCGCTCTTGTTG 39348

Qy    643 cccaggctggagtgcaagtggcgcaatcttggctcactgcaacctctgcctcccgggttca 702
           |||||||||| ||||||| |||||||||| |||||||||||||||| || |||||||||||
Db  39347 CCCAGGCTGGGGTGCAAATGGTGCAATCGGGGATTCGGGCTCACTGCAACCTTCCTGGGTTCA 39288

Qy    703 agtgattctcctgcccagcctccgagtagctgggactacaggcgcccacacacgcctag 762
           ||  |||||||||| |  ||||  |||||||||| |||||||||| ||| |||| ||
Db  39287 AGCAAATTCTCCTGCCTCAACCCCCAAGTAGCTGGGATTACAGGCGTGCCATCACACC 39228

Qy    763 ctaatttttttgtatttttagtagagatgggg-ttcacacatgttcgccaggttgatcttg 821
           ||||||| || |||| || ||||||| |||||| |||| ||||| ||||||| || |||| |
Db  39227 TGGCTAATTTTGTATTGTTACTAGAGATGGGGTTTCACTATGTTGGCCAGGCTGGTCTTCA 39168

Qy    822 atctctggacct--tgtgatctgcctcggcctcccaaagtgctgggattacaggcg 879
           | ||||| |||||  | |||  |||||| ||||||| |||||||||||||||||| |
Db  39167 AACTCCTGACCTCAGGAGATCCTCCCTGCCTCGGCCTCCCAGAGTGCTGGGATTACAGGTG 39108

Qy    880 tgagccaccacgcccggcctt--attttaattttttgttttgaaatggaatctcactc 937
           ||||||||    |||  |||| | |||| ||||||||| ||||  | |||| |||| ||
Db  39107 TGAGCCACTGTGTCCTGGCCTCAATTTTTTATTTTTTGTTTTTTGGAGACAAAGTCTCGC 39048

Qy    938 tgttacccaaggctggagtgcaatggcgcaaatctcgctcactgcctcccgg 997
           |  | |||||| |||||||| |||| || ||| |||||||||||| ||  ||
Db  39047 TCTTTCCCAGGCTGGAGTGGAAATGGCACCACCTCGGCTCACTGCACCTTCGCCTTCGG 38988

Qy    998 gctcaaggcgattctcctgtctcagcctcccaagcagctgggactggcacctgcacc 1057
           ||||||| |||||||||| | || |||||||||| ||||| | || | |||||||||
Db  38987 GTTCAAAGGCGATTCTCCTCAGCTGCTCCCAAGTAGCTACAGGCATCGTGCCACC 38928

Qy   1058 acacccgctaattttttgtattttcattagaggacgggggtttcaccacatatttgtcaggcty 1117
           | |||| |||||||| ||||||| | ||ga ||| | ||||||||||| | |||||||| |||
Db  38927 ACGCCCGGGCTAATTTTTGTATTTTAGTAGAGACGCGGGGTTTCACCATGTTGGCCAGACTG 38868

Qy   1118 gtctcaaactcctgaccctcaaggtgaccaacctgcctcccaaagtgctgtggatta 1177
           |||| ||||||||  || |||||||| || |  ||||||| ||||||||| ||||||
Db  38867 GTCTTGAACTCCTCAACTTCAGGTGATCACCCGCCTCGGCCTCCCAAAGTGCTGGGATTA 38808

Qy   1178 caggcgtgagacacctcacccagcc---ggctaatttagataaaaaaaatatgtagcaat 1233
           ||||| ||| ||||  ||   ||||   | ||||||| |||||||||||| ||| ||||
Db  38807 CAGGCATGACGCACCACCACGCTGGCCTCTCGCCCACTAATTAAAAAAAAATTTTTGTAGAG 38748

Qy   1234 gggggtcttgctatgttgccaggctggtctccaaacttctgtgcttcatgcaatccttcc 1293
```

RESULT 15

```
LOCUS        HSU63721/c    67046 bp    DNA    linear    PRI 22-NOV-2000
DEFINITION   Human elastin (ELN) gene, partial cds, and LIM-kinase (LIMK1) gene,
             complete cds.
ACCESSION    U63721
VERSION      U63721.1  GI:1657753
KEYWORDS     Williams syndrome.
SOURCE       human.
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 67046)
  AUTHORS    Osborne,L.R., Martindale,D.W., Scherer,S.W., Shi,X.-M.,
             Huizenga,J., Heng,H.H.Q., Costa,T., Pober,B., Lew,L., Brinkman,J.,
             Rommens,J., Koop,B.F. and Tsui,L.-C.
  TITLE      Identification of genes from a 500-kb region at 7q11.23 that is
             commonly deleted in williams syndrome patients
  JOURNAL    Genomics 36 (2), 328-336 (1996)
  MEDLINE    9641691
REFERENCE    2  (bases 1 to 67046)
  AUTHORS    Martindale,D.W., Wilson,M.D., Wang,D., Burke,R.D., Chen,X.,
             Duronio,V. and Koop,B.F.
  TITLE      Comparative genomic sequence analysis of the williams syndrome
             region (LIMK1-RFC2) of human chromosome 7q11.23
  JOURNAL    Mamm. Genome 11 (10), 890-898 (2000)
  MEDLINE    20458868
   PUBMED    11003705
REFERENCE    3  (bases 1 to 67046)
  AUTHORS    Martindale,D.W., Osborne,L.R., Scherer,S.W., Tsui,L.-C. and
             Koop,B.F.
  TITLE      Human genomic sequence containing the complete LIM-kinase gene
             (LIMK1) and the 3'end of elastin (ELN)
  JOURNAL    Unpublished
REFERENCE    4  (bases 1 to 67046)
  AUTHORS    Martindale,D.W., Osborne,L.R., Scherer,S.W., Tsui,L.-C. and
             Koop,B.F.
  JOURNAL    Direct Submission
             Submitted (12-JUL-1996) Biology, University of Victoria, PO Box
             1700, Victoria, BC V8W 2Y2, Canada
FEATURES             Location/Qualifiers
     source          1..67046
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7q11.23"
                     /clone_lib="chromosome 7 specific cosmid library LL07NC01
                     of Lawrence Livermore National Laboratory"
                     /note="obtained from cosmids 34b3 (complete) and 152a8
                     (partial)"
     mRNA            join(<3..182,275..319,511..636,1236..1274,3288..3359,
                     3445..3504,3756..3830,5829..5867,6079..6132,6867..6911,
                     8791..>8834)
                     /gene="ELN"
                     /note="determined using HUMELASP and Indik et al, PNAS
                     84,5680-5684,1987"
                     /product="elastin"
     gene            3..10021
                     /gene="ELN"
     CDS             join(<3..182,275..319,511..636,1236..1274,3288..3359,
                     3445..3504,3756..3830,5829..5867,6079..6132,6867..6911,
                     8791..8834)
                     /gene="ELN"
                     /codon_start=3
                     /product="elastin"
```

```
                    /protein_id="AAC13884.1"
                    /db_xref="GI:1657754"
                    /translation="LLNLAGLVPGVGVAPGVGVAPGVGVAPGVGVAPGVGVA
                    PGVGVAPCIGPGGVAAAAKSRAKVAAKAQLRAAAGLGAGIPGLGVGVGVPGLGVGAGV
                    PGLGVGAGVPGFGAVPGGALAAAKAARYGAAVPGVLGGLGALGGVGLPGGVVGAGPAAA
                    AAAAKAAAAAQFGLVGAAPGFGLVGGLGVGGLGVPGVGVGGLGVPGVGGLGGIPPAAAKAKYGAAGLGG
                    VLGGAGQFPLLGGVAARPGFGLSPIFPGGACLGKACGCRKRK"
repeat_region       1370..1798
                    /rpt_family="ALU"
                    /note="91% similarity"
repeat_region       2472..2596
                    /rpt_family="MIR"
                    /note="70% similarity"
repeat_region       complement(2717..2985)
                    /rpt_family="ALU"
                    /note="91% similarity"
repeat_region       3027..3139
                    /rpt_family="ALU"
                    /note="81% similarity"
repeat_region       3964..4251
                    /rpt_family="ALU"
                    /note="83% similarity"
repeat_region       4282..4737
                    /rpt_family="ALU"
                    /note="92% similarity"
repeat_region       5054..5370
                    /rpt_family="ALU"
                    /note="81% similarity"
repeat_region       5479..5766
                    /rpt_family="ALU"
                    /note="90% similarity"
repeat_region       6424..6708
                    /rpt_family="ALU"
                    /note="91% similarity"
repeat_region       7150..7912
                    /rpt_family="ALU"
                    /note="90% similarity"
repeat_region       complement(7408..7692)
                    /rpt_family="ALU"
                    /note="93% similarity"
repeat_region       7698..8085
                    /rpt_family="ALU"
                    /note="92% similarity"
repeat_region       9786..9791
                    /rpt_family="ALU"
                    /note="79% similarity"
repeat_region       10016..10021
                    /gene="ELN"
repeat_region       10306..10445
                    /gene="ELN"
repeat_region       complement(11671..11782)
                    /rpt_family="ALU"
                    /note="78% similarity"
repeat_region       11829..12107
                    /rpt_family="ALU"
                    /note="79% similarity"
repeat_region       12545..12817
                    /rpt_family="ALU"
                    /note="97% similarity"
repeat_region       complement(12129..12193)
                    /rpt_family="ALU"
                    /note="92% similarity"
repeat_region       13645..13747
                    /rpt_family="ALU"
                    /note="88% similarity"
repeat_region       complement(12971..13260)
                    /rpt_family="ALU"
                    /note="88% similarity"
repeat_region       14274..14554
                    /rpt_family="ALU"
                    /note="87% similarity"
repeat_region       complement(13841..14132)
                    /rpt_family="ALU"
                    /note="87% similarity"
                    /note="88% similarity"
```

```
repeat_region       14656..14736
                    /rpt_family="ALU"
                    /note="91% similarity"
repeat_region       14783..14929
                    /rpt_family="MLT1"
                    /note="85% similarity"
repeat_region       15590..15878
                    /rpt_family="L1"
                    /note="85% similarity"
repeat_region       16173..16458
                    /rpt_family="ALU"
                    /note="92% similarity"
repeat_region       16791..16995
                    /rpt_family="ALU"
                    /note="69% similarity"
repeat_region       17035..17809
                    /rpt_family="L1"
                    /note="90% similarity"
repeat_region       17922..18035
                    /rpt_family="MLT1"
                    /note="82% similarity"
repeat_region       18261..18538
                    /rpt_family="ALU"
                    /note="88% similarity"
repeat_region       complement(18545..18714)
                    /rpt_family="ALU"
                    /note="82% similarity"
repeat_region       19070..19194
                    /rpt_family="ALU"
                    /note="76% similarity"
repeat_region       19696..19989
                    /rpt_family="ALU"
                    /note="88% similarity"
repeat_region       20022..20317
                    /rpt_family="ALU"
                    /note="79% similarity"
repeat_region       20351..20626
                    /rpt_family="ALU"
                    /note="88% similarity"
repeat_region       complement(20823..21424)
                    /rpt_family="ALU"
                    /note="91% similarity"
repeat_region       complement(22002..22167)
                    /rpt_family="ALU"
                    /note="79% similarity"
repeat_region       22188..22482
                    /rpt_family="ALU"
                    /note="89% similarity"
mRNA                join(23511..23730,25433..25529,36304..36442,36762,36871,
                    38711..38917,45557..45662,45759..45925,46692..46875,
                    47552..47638,49586..48717,51327..51386,51612..51677,
                    55481..55637,60244..60299,60553..60710,60800..62185)
                    /gene="LIMK1"
                    /note="determined using HUMLIMKA, Mizuno et al., Oncogene
                    9,1605-1612,1994"
                    /product="LIM-kinase"
exon                23511..23730
                    /gene="LIMK1"
                    /number=1
gene                23511..62185
                    /gene="LIMK1"
```

```
Qy            4 tttttttgagatggagtttcgcctcttgttcgcccaggctggagtgcaatggcgcaatct 63
                |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        63627 TTTTGTTTGAGATGGAGTTTT-GCTCTTGTTGTGCCCAGGCTGGAGTGCTGAAGTGGCACGGATCT 63569
```

Query Match   40.8%; Score 587.8; DB 9; Length 67046;
Best Local Similarity 70.6%; Pred. No. 3e-164;
Matches 940; Conservative 0; Mismatches 347; Indels 44; Gaps 10;

```
Qy     64  cagctcaccgcaacctccgcgctcccgggttcaaagcggattctcctgcctcagcctcccag   123
           ||||||||| |||| |||||||| |||| ||||||| ||| ||||| |  | |||||||| ||
Db  63568  CAGCTCACCACAACTTCCGCGTCCCAGGTTCAAGCAATTCTCCTACCTTCAGCCTCCGAG    63509

Qy    124  tagctgggpattacaggcgatgtgcaccoacggctgctaattttttgtattttttagtag    183
           ||||||||| ||||||||| ||||||||| || |||||||||| |||||||     |||
Db  63508  TAGCTGGGATTACAGGCATGTGCACCACCGCTGGTTAATTTTTGTATTTTT---AGT     63453

Qy    184  agatggagttcttccatgttggtcaggctggtctcgaactcccgacctcagatgatccct   243
           ||| ||  |||||||||||||||||| ||||||| |||| |||||||| |||||||||| |
Db  63452  AGACGGGGTTTCTTCCATGTTGGTCAGGTTGGTCTTGAACCCCGACCTTCAGATGATCCGC   63393

Qy    244  ccgtctcggcctcccaaagtgctagatacagacggcgccacatgccggctctgctgg      303
           || ||| ||||| |||||||||| ||| ||||| |  |||| ||  |    |
Db  63392  CCACCTTGGCCCTTCCCAAAGTGCT-GGGATTACAGGTGTGAGTTTTTACAAGAAAACATTTA  63334

Qy    304  ctaattttttggtagaaacagggtttcactgatgttgcccaagctggtctcctgagctca   363
           |  ||||   | ||||| |  ||| ||||       | ||||||||| ||   |||| ||
Db  63333  AGTATGAAAGAGTACCTTCCGGAATATGCACGGCCCTTCCAAACCCAGCCTCTGAAGAGACTC  63274

Qy    364  agcagtccaactgcctcaagcctcccaaagtgctgggatta------------          403
           || |||| |  ||||||| || ||| ||  |||  ||
Db  63273  TTCACTTGTCCTGGCCCACCAACCTTCCAGCCTCCACCTAAAAACTGGGCTTGTTCTTCTT   63214

Qy    404  caggcgtgcagccgccgtggcctggcctttttattttttattttttaagacacagtgtcccac   463
           |||| | ||| ||  |   ||||| |  |||||||||||||||||| ||| |||| | ||
Db  63213  CCCCCCTTCCATCCCTGCCTTTTTTTTTTTTTTTTTTTTTTTTGAGACAGAGTCTTGTTT    63154

Qy    464  tcttacccaggatgaagtgcagtggtgtgatcacagctcactgcagccttcaactcctga   523
           |  |||||||| | ||||||||||| | | |||||||||||||||| || |||| |||
Db  63153  TGTTGCCCAGGCTGGAGTGCAGTGGCGCGAAACATGGCTCACTGCAGCCTCCGCCTCCCGG   63094

Qy    524  gatc-aagcaatcctcctgcctcagcctcccaagtagctgggactacaggcgcccacact   582
           || | ||| ||||||| | |||| ||||||||||||||||| | |||| ||||||||  ||
Db  63093  GCTCAAAGGAATCCTCCCAACCCCCACGTAGCTGGGATTTACAGGCGCCCATCACACC      63034

Qy    583  acacctggctaattttttattttttatttttttggagacagagtctcaactctgtca      642
           |  | | |  ||||||||||||||   ||||| ||| | |||| ||||| | |||
Db  63033  ATGCCCAACTAATTTTTTTTTTTTT----TTTTTGAGATGGAGTTTCGCTCTTGTTG       62982

Qy    643  cccaggctggagtcgcagtggcgcaatcttggctcactgcaacctctgcctcccgggttca   702
           |||||||||| || ||||||| |||| | ||||||||||||||||||||||||||||||||
Db  62981  CCCAGGCTGGGGTGCAATGGTGCGATCTCGGCTCACTGCAACCTCTGCCTCCCGGGTTCA    62922

Qy    703  agttattctcctgccccagcctcctgagtagctgggactacaggcgcccaccagcctag    762
           ||  ||||||| | |||| || || ||||||||||| | ||||  |||| ||| || |
Db  62921  AGCAATTCTCCTCGCCTCAACCCCCCAGTAGCTGGGATTTACAGCATGGCCATCACACC     62862

Qy    763  ctaattttttgttatttttagtagagatggggttcaccatgttgccaggttgatctg      821
           ||||||| ||||||| |||||||||||||||||| ||| |||| |||| ||| |||
Db  62861  TGGCTAATTTGTTATTGTTAGTAGAGATGGGGTTTCACTATGTCAGGCTGGTCTCA       62802

Qy    822  atctctggacct--tgtgatctgcctgcctcggcctcccaaagtgctgggattacagggg   879
           | |||| |||||  |  | || ||| |||||||||||||||||||||||||||||||||
Db  62801  AACTCCTGACCTCAGGAGATCCTCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGGTG   62742

Qy    880  tgagccacacgcccggcctt--atttttaattttgttgttgaaatggaatctcactc     937
           |||||||| | ||||| |||  ||||||  ||||| |||| | | |||| ||| |
Db  62741  TGAGCCACTGTGCCCGGCCTAATTTTTTTTTTTTTTTTTGGAGACAAAGTCTCGC        62682

Qy    938  tgttacccaggctgggagtgcaaatggcaaatctcggctcaactctgctccggg          997
           | ||| ||||||||||||||| ||||||||  | |||||||| |||||| | |||
Db  62681  TCTTTCCCAGGCTGGGAGTGGAAATGGCACCACCTCGGCTCACTGCAACCTCTGCCTTTCGG  62622

Qy    998  gctcaagcgaattcctgtctcagcctcccaagcaagctgggattacgggcgaacctgccacc  1057
           | ||||||| | ||||| |||||||||||||| |||| |||  |||||  |||| || ||||
Db  62621  GTTCAAGCGACATTCCTCTGCCTCAGCCTCCCAAGTAGCTACAGGGCCATGTGCCACC      62562

Qy   1058  acacccgctaattttttgtatttttcattagatgggggtttcaccatatttgtcaggctg  1117
           || ||||| |||||||| |||||||  || || || ||| |||||||| ||| ||| ||
Db  62561  ACGCCCGGCTAATTTTTGTATTTTTTAGTAGAGACGAGTTTTCACCATGTTGGCCAGACTG   62502

Qy   1118  gtctcaaactcctgacctcaggtgacccacctccaaagtgctgggatta              1177
           |||||||||||||||||||||||||  ||||| ||||||||||||||||
Db  62501  GTCTTGAACTCCTAACTTCAGGTGATCCACCCGCCTCGGCCTCCCAAAGTGCTGGGATTA   62442

Qy   1178  caggcgtgagccacctcacccagcc----ggctaattttagataaaaaaatatgtagcaat  1233
           ||||| |||||||| |||||| |  |    | | |||||| |||||||||| || ||||||
Db  62441  CAGGCATGAGCCACCACCACCACCTGGCCTCTGCCCAACTAATTAAAAAAATTTTGTAGAG   62382

Qy   1234  gggggttcttgctatgttgcccaggctggtcttcaaaccttctggcttcatgcaatccttcc  1293
           |||| ||||| |||||||| |||| ||||||||||| |||||||||| |  ||||| | | ||
Db  62381  ACAGGATCTCACTATGTTGTCCAGACTGGTCTTGAACTCCTGGGCTTAAGGATCGTCCT     62322

Qy   1294  aaatgagccac  1304
           ||| ||||| |
Db  62321  GTCTCAGCCTC  62311
```

Search completed: April 24, 2002, 19:40:35
Job time: 9314 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 24, 2002, 17:03:15 ; Search time 1503.67 Seconds
                     (without alignments)
                     12943.405 Million cell updates/sec

Title: US-09-964-678A-1
Perfect score: 1442
Sequence: 1 ttttttttttgagatggag.............ttaaacaaagcttagagca 1442
Scoring table: IDENTITY_NUC
                     Gapop 10.0 , Gapext 1.0
Searched: 13736207 seqs, 6748477542 residues
Total number of hits satisfying chosen parameters: 27472414

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Database :   EST:*
                     1: em_estba:*
                     2: em_esthum:*
                     3: em_estin:*
                     4: em_estmu:*
                     5: em_estov:*
                     6: em_estpl:*
                     7: em_estro:*
                     8: em_htc:*
                     9: em_est1:*
                     10: gb_est2:*
                     11: gb_htc:*
                     12: gb_gss:*
                     13: em_gss_hum:*
                     14: em_gss_inv:*
                     15: em_gss_pln:*
                     16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB ID	Description
1		396.2	27.5	1842	11	BC011119	BC011119 Homo sapi
2	c	391.4	27.1	877	12	AQ739838	AQ739838 HS_5505_A
3	c	383.6	26.6	1345	9	AV762220	AV762220 AV762220
4	c	378.4	26.2	767	9	AV700498	AV700498 AV700498
5		373.8	25.9	658	12	AQ393450	AQ393450 CITBI-E1-
6		373	25.9	911	12	AQ746594	AQ746594 HS_2278_A
7	c	368.8	25.6	970	10	BM468547	BM468547 AGENCOURT
8		367.6	25.5	5556	12	AQ839814	AQ839814 260113-C5
9		366.2	25.4	839	9	AV700988	AV700988 AV700988
10		360	25.0	1148	9	AV761207	AV761207 AV761207
11	c	359.8	25.0	667	12	AG150197	AG150197 Pan trogl
12	c	358.8	24.9	687	12	AG186062	AG186062 Pan trogl
13	c	357.8	24.8	679	10	BF346320	BF346320 602018525
14	c	357.4	24.8	722	12	AQ583981	AQ583981 RPCI-11-4
15	c	356.6	24.7	1259	11	BC009590	BC009590 Homo sapi
16	c	355.4	24.6	1808	11	BC017826	BC017826 Homo sapi
17	c	355	24.6	740	12	AQ878530	AQ878530 HS_3081_B
18	c	353.2	24.5	689	12	AQ318783	AQ318783 RPCI11-92
19		350.6	24.3	723	9	AL601019	AL601019 DKFZp313F
20	c	350	24.3	737	9	AV700545	AV700545 AV700545
21	c	346.6	24.0	1875	11	BC009270	BC009270 Homo sapi
22	c	345.6	24.0	1771	11	BC018833	BC018833 Homo sapi
23		345.2	23.9	749	9	AV699709	AV699709 AV699709
24	c	344.8	23.9	792	10	BI225234	BI225234 602949941
25	c	343.6	23.8	738	10	BG546184	BG546184 602573405
26	c	343.6	23.8	758	10	BG563032	BG563032 602581816
27		342.8	23.8	617	10	BI861844	BI861844 603388872
28	c	341.2	23.7	714	10	BG289947	BG289947 602381363
29		341	23.6	634	12	AQ543763	AQ543763 RPCI-11-3
30	c	340.6	23.6	648	9	AV762783	AV762783 AV762783
31	c	340.4	23.6	621	12	AQ636660	AQ636660 RPCI-11-4
32		339.6	23.6	1915	11	BC012949	BC012949 Homo sapi
33	c	338.4	23.5	660	12	AQ381918	AQ381918 RPCI11-13
34	c	338.2	23.4	607	12	AQ583978	AQ583978 RPCI-11-4
35	c	337.6	23.4	691	10	BG392240	BG392240 602410376
36	c	337.2	23.3	942	12	AQ743787	AQ743787 HS_5501_A
37		335.6	23.3	839	10	BG164166	BG164166 50243050
38		335	23.2	580	12	AQ389063	AQ389063 RPCI11-13
39	c	335	23.2	616	10	BG028665	BG028665 60292869
40	c	334.8	23.2	839	9	AV756726	AV756726 AV756726
41	c	334.6	23.2	990	9	AW600804	AW600804 N26-20 Hu
42		334	23.2	615	12	AG144700	AG144700 Pan trogl
43	c	334	23.2	688	9	AV762145	AV762145 AV762145
44		333.6	23.1	973	10	BF968610	BF968610 60271031
45		333.2	23.1	5528	12	AB012089	AB012089 Homo sapi

RESULT   1
ALIGNMENTS

LOCUS   BC011119   1842 bp   mRNA   linear   HTC 30-JUL-2001
DEFINITION   Homo sapiens, clone IMAGE:3047997, mRNA.
ACCESSION   BC011119
VERSION   BC011119.1 GI:15029795
KEYWORDS   HTC.
SOURCE   human.
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1842)
  AUTHORS   Strausberg,R.
  TITLE   Direct Submission
  JOURNAL   Submitted (25-JUL-2001) National Institutes of Health, Mammalian
             Gene Collection (MGC), Cancer Genomics Office, National Cancer
             Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
             USA
COMMENT   NIH-MGC Project URL: http://mgc.nci.nih.gov
             Contact: MGC help desk
             Email: cgapbs-r@mail.nih.gov
             Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Baylor College of Medicine Human Genome
             Sequencing Center
             Center code: BCM-HGSC
             Web site: http://www.hgsc.bcm.tmc.edu/cdna/
             Contact: villalon@bcm.tmc.edu
             Villalon, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia,
             A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,
             Muzny,D.M., Gibbs,R.A.
             Clone distribution: MGC clone distribution information can be found
             through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
             Series: IRAK Plate: 19 Row: o Column: 1
             This clone has the following problem: incomplete processing.
FEATURES   Location/Qualifiers
     source   1..1842
                     /organism="Homo sapiens"

```
                        /db_xref="taxon:9606"
                        /clone="IMAGE:3047997"
                        /tissue_type="Cervix, carcinoma"
                        /clone_lib="NIH_MGC_12"
                        /lab_host="DH10B"
                        /note="Vector: pCMV-SPORT6"

Qy  Query Match
    Best Local Similarity   27.5%;   Score 396.2; DB 11; Length 1842;
    Matches 746;  Conservative   0;  Pred. No. 1.8e-77;
                                     Mismatches 263;   Indels  52;  Gaps  14;

BASE COUNT      371 a    522 c    475 g    474 t
ORIGIN

Qy       2 tttttttttgagatggagggagtttcgctcttgttgccaaggctgaggtgcaatggcgcaat   61
           ||||||||| |||||  ||||| ||||||| |  ||||  | ||| ||||||  ||||||||
Db     570 TTTTATTTGTTTTGAGATGGAGTCCGCGGCTTGTGTCGCCCACGCTGGAGTGCAGTGGCGTGATC  629

Qy      62 ctcagctcaccgcaacctccgcctcccgggttcaagcgattctcctggctcagcctccc     121
           |||||  || ||||| ||| |||||||| ||||||||  ||||| || |||||||||||
Db     630 CTCGGCTCACTGCAAGCTCCGCCTCCCGGGTTCACACCTTTCTCCTGGCTCAGCCTCCCA     689

Qy     122 agtagctcgggattacaggcgatgtgcacccacgctcggtctaatttgtatttttttagt  181
           ||||||||||||  | ||  |||||||||||||||||||||||  || ||||||| ||||
Db     690 AGTAGCTGGGGACTACTGGTGCCCGCCACCACGCCCGGCTAATTTT-TTTCTATTTTTAGT  748

Qy     182 agagatggagtttctccatgttggtcaagctggtctgttctccgaactccgacctcagatgatcc  241
           |||| |||||||||| |||| |||||||||||||||||| ||||| |||| |||||||| |
Db     749 AGAGACGGGGTTTCACCGTGTTACCCAGGATGGTCTCGATCTCCTGACCTCGTGATGATCC  806

Qy     242 ctccgtctcggcctcccaaagtgctagatacaggcgtgccacacaatgcccggctctgcct  301
           |||  | ||||||  ||||||||| |||| |||| || ||||| ||||||| |||||||||
Db     807 ACCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCCCGGCCTTA  866

Qy     302 ggctaatttttgtggtagaaacagggtttcactggttgtcccaagctg-gtcctctgagc  360
           || ||||||||||| || || |||||||||||||| || ||| |||| ||||||| |
Db     867 TTTTATTTTTTAATAGAGACGGGGTCTCACTATGTTGCCCAGGCTGGTCTCAAACTCCTGGGC  926

Qy     361 tcaagcaagtccacctgcctcagcctcccaaagtgctgggattacag--gcgtgcagccgt  418
           ||||| || |||| || ||||||| ||||||||||||| ||||||| |  |||||| |
Db     927 TCAAGCCAATCCTCCCACTT-GGCCTCCCAAAGTGCTGGGATTACAGTTGTGAGCCACTGT  985

Qy     419 gcctggccttttttttttttttttaagacacaggtgtccacactcttaccaggtatga  478
           |||||||  |||||||||||||| || | ||| |||||| || ||| ||||| ||| |
Db     986 GCCTGGCCTCTTTTTTTTTTTTTTTGAGATAGA------AATCTCCCTGTCACCCAGGTTGG  1038

Qy     479 agtgcagtggtgtgatcaacagctcact-----------gcagccttcaac  517
           |||||| |||  || |||  |||||||             |||||||||||
Db    1039 ACTGCAGTGGTGCGATCTCAGCTCACTGCAACCTCCGGCTCAGTTGCCTTCAAC  1098

Qy     518 tcctgagatcaagc-atcctcctgcctcagcctcccaaaagacatgc  576
           ||||||| |||||| |||||||||||||||||| |||||  |||||
Db    1099 TCCTGGGTTCAAGCAAATTCTCCTGCCTCAGCCTGAGATTACAGGCGCAT  1158

Qy     577 accactacactggctaattttttattttttgagacagagtctcaact  636
           || |||||| |||||| |||||||||||||||||||||| ||| |||
Db    1159 GCCACCACGCCTGGCTAATTTTTGTGTTTTTTTTTGTTTTGAGACAGAGTCTCT----CT  1215

Qy     637 ctgtcacccaggctggagtcgcagtggcgcgaatcctcgcagcctctgcctccg  696
           ||||||||||||||||||| |||| |||||| ||||| ||||| |||||| ||||
Db    1216 CTGTCA--CCAGGCTGGAGTGCAGTGGCGTGATCTGGCTCACTGCAACCTCCGCCTCCCG  1274

Qy     697 ggttcaagtattctcctgcccccagcgcctcctgagtagctgggactacaggcgcccaccac  756
           |||||||| | ||||| | | ||||||||||||||| |||||||| |||||| |||||||||
Db    1275 GGTTCAAGCGATTCTCCTGCC-GCAGCCTCTCAAGTAGCTGGGACTACAGGCGCCCACCAC  1333

Qy     757 gcctagctaat-tttttgttatttttagtagagacggggg-ttcaccaatgttcgccaggtt  814
           ||||| ||||| ||||| ||||| ||||| ||| |||||| ||||||  |||||||||||
Db    1334 CATGCCCGGCTAATTTTGTGTATTTTTAGTAGAAGAGTGGGGTTTCACCATGTTGGCCAGGAT  1393

Qy     815 gatcttgatctctggaaccttgtgatctcgccgtctccccaaagtgctgggattac  874
           ||| ||| |||||||||||| ||||||||||| ||||| |||||||||||||||||
Db    1394 GGTCTCAAACTCCTGACCTCATGATCACCACCTCGCTCCCAAAACTGCTGGGATTAC  1453
```

```
RESULT  2

LOCUS       AQ739838/c
DEFINITION  AQ739838   877 bp   DNA   linear   GSS 16-JUL-1999
            HS_5505_A1_C09_T7A RPCI-11 Human Male BAC Library Homo sapiens
            genomic clone Plate=1081 Col=17 Row=E, DNA sequence.
ACCESSION   AQ739838
VERSION     AQ739838.1     GI:5517360
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 877)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
   MEDLINE   99380589
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieterdejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
            or from Resear h Genetics (info@resgen.com). BAC end Web Server:
            http://www.htsc.washington.edu
            Plate: 1081 row: E column: 17
            Seq primer: T7
            Class: BAC ends
            High quality sequence stop: 877.
FEATURES            Location/Qualifiers
     source         1..877
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /clone="Plate=1081 Col=17 Row=E"
                    /clone_lib="RPCI-11 Human Male BAC Library"
                    /sex="male"
                    /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI.
                    Male blood DNA was isolated from one randomly chosen donor
                    and partially digested with a combination of EcoRI and
                    EcoRI Methylase. Size selected DNA was cloned into the
                    pBACe3.6 vector at EcoRI sites"
BASE COUNT      181 a    242 c    203 g    251 t
ORIGIN
```

```
Qy     875 aggccgtgagtccacacgcccggcgttatttttgttgtttggaatctcca  934
           ||| |||  |||||||||||||| |||| |||  |||||||||| |||||
Db    1454 AGGTGTGAGCCACATCACGCCCAGCGCCTTTTTTTTTTTTTTTTTT-----TAGATCTACC  1504

Qy     935 ctctgttacccaggctggagtgcaatgggcaaatcctcggctcactgcaactctgcctcc  994
           ||||| |||||||||| ||||| | | |  ||  |||||||||||||| ||||||||||
Db    1505 CTCTTGTTGTTCCCCAGGGGGAGTTCAGTGGGCGTGATCTGGCTCACTGCAACTGCCTCC  1564

Qy     995 cgggctcaagcgagttctcctgtctcagcctcccaagcagct  1035
           ||||||||||||| ||||||||| || |||||||||||| |
Db    1565 CGGGTTCAAGCGAATTCTCCTGCCTCAGCCTCCCGAGTGTCT  1605
```

```
Qy  Query Match
    Best Local Similarity   27.1%;   Score 391.4; DB 12; Length 877;
    Matches 514;  Conservative   0;  Pred. No. 2.3e-76;
                                     Mismatches 126;  Indels   7;  Gaps  4;

BASE COUNT      251 a    181 c    242 g    203 t
ORIGIN

Qy     568 aagacatgcacactacacctggctaattttttattttttgagacaga  627
           || || ||||   || ||||||||||||||||||||||||||||||
```

RESULT 3

LOCUS   AV762220   1345 bp   mRNA   linear   EST 19-OCT-2000
DEFINITION   AV762220 MDS Homo sapiens cDNA clone MDSCBD07 5', mRNA sequence.
ACCESSION   AV762220
VERSION   AV762220.1 GI:10920068
KEYWORDS   EST.
SOURCE   human.
ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1345)
AUTHORS   Gu,J., Zhao,M., Huang,Q., Xu,X., Li,Y., Peng,Y., Song,H., Xiao,H.,
             Gu,Y., Li,N., Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng
             ,L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G.,
             Yang,Y., Gao,G., Zhang,Q., Chen,S., Han,Z. and Chen,Z.
SOURCE   Homo sapiens
ORGANISM   Homo sapiens cDNA MDS clones
KEYWORDS   EST.
COMMENT   Contact: Zeguang Han
             Chinese National Human Genome Center at Shanghai
             351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
             201203, P. R. China
             Tel: 86-21-50801919(ex.45)
             Fax: 86-21-5801922
             Email: hanzg@chgc.sh.cn
             This clone is available at CHGC in Shanghai.
FEATURES   Location/Qualifiers
     source   1..1345
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"

BASE COUNT   388 a   316 c   354 g   277 t   10 others
ORIGIN

Query Match   26.6%;   Score 383.6; DB 9; Length 1345;
Best Local Similarity 82.8%; Pred. No. 1.1e-74;
Matches 502; Conservative   0; Mismatches   86; Indels   18; Gaps   5;

```
Qy  680 AACACCTCCCCAGTAATGAATGTCTTTTGTTTTTGTTTTTTTTGTTTTTTTGTTTTTTTGAGACAGA 621
Qy  628 gttcaacctgtcaccaggctggagtgcagtggcgcaatcttggctcaactggcaacctc 687
Db  620 GTCTC-ACTTTGATGCCCAGGCTGGAGTACAGTGGTGCGAATCTCAGCTCACTGCAAACCTT 562
Qy  688 tgcctcccggggttcaagttattctcctgcccccagcctcctgagtagctgggactacaggc 747
Db  501 GCCTCCCACCACGCCCAGCTAATTTTTATATTTTTAGTAGAGACA-GGGTTTCACCATGT 443
Qy  805 tcgccaggttgatcttgatctctgggacct--tgtgatctgcctgctcggcctcccaaag 862
Db  442 TGGGCAGGGCTTGTCTTGTTGCTGTTATTT...403
Qy  863 tgctgggattacaggcgtgagcacacgggcccggggcttattttttttgtttgttttga 922
Db  382 TGCTGGGATTATAAGCGTGAGCCACCGCACCAACCAGTAATTTTTTTTTTTTTGTTTTTTTGA 323
Qy  923 aatggaatctcacttgttacccaaggctggagtgcaaatgcccaaatctccggctcactgca 982
Db  322 GATGGAGTCTCACTCTGTTACCCAGGCTGGAGTGCAGTGGCGCGATCTTGGCTCACCGCA 263
Qy  983 accttgcctcccgggctccaagcgattctctgcctcccaagcagctggatta 1042
Db  262 ACCTCTGCCTCCCAGGTTCAAGCGATTCTCCTGCCTCAGCCTGCCGAGTAGCTGGGAATTA 203
Qy 1043 cgggcacctgccacacaccccgctaatttttgtatttttcattagagacggggtttcacc 1102
Db  202 CGGGCACGTGCCACCATGCCCGGCTAATTTTTTGTATTTTTTAGTAGAGACAGGGGTTTGCGC 143
Qy 1103 atatttgtcaagctggtcttcaaactcctgacctccaagtgatcccacctgcctcaagcctcc 1162
Db  142 AtGTTGTGTCCAGGGTGGTGTCTCGAACTCCTCACTCAAGGTGAGCCGAGGATCCTC 83
Qy 1163 aaagtgctgggattacaggcgtgagcaacctcacccacgccggctaat 1209
Db   82 AAAGTGCTGGGATTATAGGCATGAGCCACCACCACATCTGGCCCAGTAAT 36
```

RESULT 4

LOCUS   AV700498   767 bp   mRNA   linear   EST 16-JAN-2002
DEFINITION   AV700498 GKC Homo sapiens cDNA clone GKCAAC05 3', mRNA sequence.
ACCESSION   AV700498
VERSION   AV700498.1 GI:10302469
KEYWORDS   EST.
SOURCE   human.
ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 767)
AUTHORS   Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,
             Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W.,

```
Qy  600 attttattttattttttttgagacagagagtctcaactctgtcacccaggctgagtgcag 659
Db  606 ATTTTTNNTNTTTTTTTTTTTTGAGACGGAGTCTC-ACTCGTGTGCCCAGGCTGGAGTGCAG 548
Qy  660 tggcgcaaatctggtcaactgcaacctctgcctcccgggttcaagttattctcctgcccc 719
Db  547 TGGCGTGATCTCAGCTCACTGCAACCTCCGCCTCCTGGGTTCAAGCAATTCTCCTGCCTC 488
Qy  720 agcctcctgagtagctgggactacaggcgc--ccacacacgcccagctaatttttgta 776
Db  487 AGCCTCCTGAGTAGCTGGGACTACAGGCGCCCGCCACCACGCCCAGCTAATTTTTGTATT 428
Qy  777 ttttttagtagagatggggttcacacatgttccgcccaggttgatctctgatccttgt 836
Db  427 TTTAGTAGAGACA-GGGTTTCATCATGTTGGCCAGGCTGGTCTTGAACTCCTGACCTCGT 369
Qy  837 gatctgcctcggcctcccaaagtgctgggattacaggcgtgagccaccacgccccgg 896
Db  368 GATCCACCCACCGGACTCCTCAAGTGCTGGGATTACAGGCGTGAGCCACCATGCCCGG 309
Qy  897 c-------ttattttaattttgttttgttttggaaatgggaatctcactctgttaccc 945
Db  308 CCTCATTTGCCAtTCTTTTTTTTTTTTTTTCTTTGAGATGGAGTCCACTCTATTGCCC 249
Qy  946 aggctgagtgcaaatgccaaatctcggctcactgcaaacctctgcctcccgggtccaagc 1005
Db  248 AGGCTGGAGTGCGTGTGGGCACAAATCTCGGCTCACTGCAAACCTCTGCCTCCCGGATTCAAGT 189
Qy 1006 gattctcctgtctcagcctcccaagcagctgggattacgggcgccacacacccccg 1065
Db  188 GATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGCATGAGCCACCACCATGCCCAG 129
Qy 1066 ctaatttttgtatttccattagagacgggggtttccaccatatttgtcaaggctcggtctctcaaa 1125
Db  128 CTAATTTTTTTGTATTTTTACTAGAGACACAGGGTTTCACCAACCATTGGTGAGGCTGGTCTCAAA 69
Qy 1126 ctcctgacctcaggtgacccaccctgcctccaaaagtgctgggattacaggcgtg 1185
Db   68 CTCCTGACCTC--GTGATCGCCCGCCTCACCCAAAGTGCTCGGGATTACAGGCGTG 11
Qy 1186 agccac 1191
Db   10 AACCAC 5
```

Left column

TITLE Insight into hepatocellular carcinogenesis at transcriptome level by comparing gene expression profiles of hepatocellular carcinoma with those of corresponding noncancerous liver

Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X., Hu,G., Gu,J., Chen,Z. and Han,Z.

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
MEDLINE 21625106
COMMENT Contact: Zeguang Han
 Chinese National Human Genome Center at Shanghai
 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
 201203, P. R. China
 Tel: 86-21-50801919(ex.45)
 Fax: 86-21-50801922
 Email: hanzg@chgc.sh.cn
 This clone is available at CHGC in Shanghai.

FEATURES Location/Qualifiers
 source 1..767
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /clone="GKCAC05"
  /clone_lib="GKC"
  /tissue_type="hepatocellular carcinoma"
  /dev_stage="Adult"
  /lab_host="SOLR"
  /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2: XhoI"

BASE COUNT 148 a 218 c 184 g 216 t 1 others
ORIGIN

Query Match 26.2%; Score 378.4; DB 9; Length 767;
Best Local Similarity 83.1%; Pred. No. 1.7e-73;
Matches 481; Conservative 0; Mismatches 87; Indels 11; Gaps 4;

```
Qy     635 ctctgtcaccaggctggagtgcagtggcgcaatcctcagctctgcctcc              694
           |||||||||||| ||| | |||||||||||||||| ||||||| |||||
Db      13 CTCTGTCACCCAGGCTGGAGTGCAGTGGCGCGATCTCGGCTCACTGCAAGCTCCTTCCTCC  72

Qy     695 cgggttcaagttattctcctgcccagcctcggagtggactacaggcgg--ccc        751
           |||||||| |||||||||||||||||||||||||||||||||||||||  |||
Db      73 CGGGTTCACCCAATTCTCCTGCCTCAACCTCCCGAGTAGCTGTGAATACAGGCGTCCGCC  132

Qy     752 accacgcctagctaattcttttgtattttgtagagatgggg-ttcaccatgttcgcca    810
           ||||||||||| ||||||||  ||||||||||||||||||| |||||| ||| ||||
Db     133 ACCACGCCCGACTAATTTTTTGTATTTTTAGTAGAGACGGGGTTTCACCGTGTTAGCCA   192

Qy     811 ggttgatcttgatctctggacctgtgatctgcctgccctcccaaagtgctggga      870
           |||||| ||| || ||||||| ||||| ||||||||||||||||||||||||||
Db     193 GGATGGTCTGGAANTCTCCTGACCTCGTGATCTGCCCGCCTCGGCCTCCCAAAGTGCTGGGA 252

Qy     871 ttacaggcgttgagccaccacgcccggcgcttatttttttgtttgaaatggaat      930
           ||||||| ||| ||||||||||||||||||  ||||| |||||||| |||||||
Db     253 TTACAGGCGTTGAGCCACCACGCCCGGCCGATTCTGCCTCGGCTCCCAAAGTGCTGGGA    306

Qy     931 ctcactctgttacccaggctggagtgcaaatctcggctcactgcaacctctgc       990
           ||||||| | |||||| ||||||||||||| |||| |||| ||||||||| |||
Db     307 CTCACACTGTTGCCCAGGCTGGAGTCGCAGTGGCTCGATCGCTCACTGCAACCTCCGC    366

Qy     991 ctcccgggttctcaaggcggttctcctgcctcagcctcccggtattacgggggcacc    1050
           |||||||| |||| | |||||| ||||| ||||||||||||| ||||| |||| || 
Db     367 CTCCTGGGTTCAAGAGATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGTGC   426

Qy    1051 tgccaccacacccgcta-atttttgtatttttcattagagcggggtttcaccatattg   1109
           |||||||||||||||| |||| | |||||||||  |||| | ||||| |||||| ||
Db     427 CACCACCACGCCTGGCTAGTTTTTGTATTTTTAGTAGAGACGGGGTTTCACCATGTTG    486

Qy    1110 tcaggctggtctcaaactcctgacctcaggtgacccacctgcctcagccttccaaagtgc  1169
           |||| |||| | ||||| |||||||||||| |||| |||  ||||| ||||| ||||||||
Db     487 CCAGGCTGGTCTCGAACTCCTGACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAGTGC   546

Qy    1170 tgggattacaggcgtgagccaccacccacccaggccgctaa                   1208
           ||||||||||||| |||||||||||||||||||||  |||
Db     547 TTGGATTACAGGCGTGAGCCACCACCATTACCTGCCAGCAAA                   585
```

Right column

RESULT 5

LOCUS AQ393450 658 bp DNA linear GSS 06-MAR-1999
DEFINITION CIT-BI-E1-2546H6.TF CIT-BI-E1 Homo sapiens genomic clone 2546H6, DNA sequence.
ACCESSION AQ393450
VERSION AQ393450.1 GI:4364473
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 658)
AUTHORS Zhao,S., Adams,M.D., Nierman,W., Malek,J., Shizuya,H., Simon,M. and Venter,J.C.
TITLE Use of BAC End Sequences from CalTech Libraries for Sequence-Ready Map Building
JOURNAL Unpublished (1997)
COMMENT Contact: Shaying Zhao, William Nierman, Mark Adams
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: hbe@tigr.org
 Clones are available from Research Genetics (info@resgen.com). BAC end search page:
 http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
 Seq primer: M13-21
 Class: BAC ends.

FEATURES Location/Qualifiers
 source 1..658
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /clone="2546H6"
  /clone_lib="CIT-BI-E1"
  /sex="male"
  /cell_type="sperm"
  /note="Vector: pBeloBAC11; Site_1: EcoRI; Site_2: EcoRI;
  CalTech Human BAC Library D"

BASE COUNT 106 a 183 c 163 g 206 t
ORIGIN

Query Match 25.9%; Score 373.8; DB 12; Length 658;
Best Local Similarity 80.9%; Pred. No. 1.8e-72;
Matches 512; Conservative 0; Mismatches 107; Indels 14; Gaps 6;

```
Qy     596 ttttatttttattttttaattttttgagacagagtccaactctgtcaccaggctggagt  655
           |||| ||||||||||||||||||||| ||||||||| || |||||||||||||||||||
Db       8 TTTTTTTTTTTTTTTAATTTTTTTTTGGAGACGGAGTCTC-ACTCTGTCACCAGGCTGGAGT 66

Qy     656 gcagtggcgcaatcttggctcactgcaacctctgcctcccgggttcaagttattctcctg  715
           |||||||| ||| || ||||||||||||||||||||||||||||||||| |||||||||
Db      67 GCAGTGGCGCGATCTCGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGCGATTCTCCTG  126

Qy     716 cccagcctcctgagtagctgggactacaggcgcccaccacccacggcctagctaat-ttttg  774
           ||||| ||||  |||||||||||||||||||| |||||||||||  ||||||| | ||||
Db     127 CCTCAGCCTCCCGAGTAGCTGGGACTAGAGGTGTGCGCCACCACGCCCGGATAATTTTTG  186

Qy     775 tattttttagtagagatgggg-ttcaccatgttcgccaggttgatcttgatctctgggacct  833
           ||||||| |||||||| |||| |||||||||| ||||||| |||| |||| |||| ||||
Db     187 TATTTTTTAGTAGAGACGGGGTTTCACCATGTTGGCCAGGCTGGTCTCGATCTCCTGGACCT  246

Qy     834 tgtgatctgcctgcctcggcctcccaaagtgctgggattacaggcgtgagccaccacgccgcc  893
           |||||||| || || ||||||||||||||||||||| |||||||||||||||||||| || 
Db     247 CGTAATTCTCCTGCCTTGGCCTCCCAGAGTGCTGGGATTACAGGCGTGAGCCACCACGCCAGCC  306

Qy     894 cggc-------ttattttttaattttttgttttgaaatctcactc-tgttac            943
           ||||       |||| |||||||| |||| ||||||||| |||| ||||||
Db     307 CAGCCCCTCGGCCCTCTTTTTTTTTTTTTTGGTTTTCCGAGGGGGAGTTTCGCTCTTGTTGC   366
```

RESULT 6

```
LOCUS       AQ746594    911 bp   DNA    linear   GSS 19-JUL-1999
DEFINITION  HS_2278_A2_B05_MR CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=2278 Col=10 Row=I, DNA sequence.
ACCESSION   AQ746594
VERSION     AQ746594.1  GI:5533752
KEYWORDS    GSS.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 911)
AUTHORS     Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., and
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
TITLE       Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
JOURNAL     Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
MEDLINE     99380589
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones may be purchased from Research Genetics (info@resgen.com).
            BAC end Web Server: http://www.htsc.washington.edu
            Plate: 2278  row: I  column: 10
            Seq primer: M13 Reverse
            Class: BAC ends
            High quality sequence stop: 911.
FEATURES             Location/Qualifiers
     source          1..911
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="plate=2278 Col=10 Row=I"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /sex="male"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"
BASE COUNT    193 a   237 c   245 g   235 t   1 others
ORIGIN

Query Match   25.9%;  Score 373;  DB 12;  Length 911;
Best Local Similarity  79.1%;  Pred. No. 2.7e-72;
Matches  493;  Conservative  0;  Mismatches 125;  Indels 5;  Gaps 4;

Qy   944 ccaggcgtggagtgcaatggccaaatctcggctcactgcaacctctgcctcccgggctcaa 1003
             |||| || ||||||| | |||  ||||||||||||||||||||||||||||||||| | ||
Db   367 CCAGGCTGGGAGTGCAGTGGTGCGAATCTCGGCTCACTGCAACCTCTGCCTCCCGGGTTCAA 426

Qy  1004 gcgattcctcgtcttcagcctcccaagcagcagcagtgggacaactgccacacacacc 1063
             | ||||||| | ||||||||||||||| |||  ||||||| |||| |||||||||||
Db   427 GTGATTCCTCCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGCGTGCACCACCACC 486

Qy  1064 cgctaa-tttttgtatttcattagaggcggggtttcaccacatattgtcaggctggtctc 1122
             ||||| |||||||||||| | |||||| ||||||||||| ||| ||| |||||||||||
Db   487 AGCTAATTTTTGTATTTTGTTTTAGTAGAGACGGGGTTTCACCATGTTGGCCAGGCTGGTCTC 546

Qy  1123 aaactcctgaccctcaggtgacccacctgcctccaaagtgctgggattacaggc 1182
             ||||||||||| |||||||||| |  |||| | || ||||||||||||||| ||
Db   547 GAACTCCTGACCTCAGGTGATCTACCCGCCTCGGCCCTCCAAAGTGCTGGGATTATAGGT 606

Qy  1183 gtgagcgaccaccacccggccgcgtaatttagat 1215
             ||||||  ||||| || |||||| ||| ||| ||
Db   607 GTGAGCCACCACACCCGGCCGCCCCATTTTGTAT 639
```

RESULT 7

```
BM468547/c
LOCUS       BM468547    970 bp    mRNA    linear   EST 05-FEB-2002
DEFINITION  AGENCOURT_6475537 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5578163
            5', mRNA sequence.
ACCESSION   BM468547
VERSION     BM468547.1  GI:18517589
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 970)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DFP
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12332  row: m  column: 12
            High quality sequence stop: 671.
FEATURES             Location/Qualifiers
     source          1..970

Db    42 TTTTGTTGTTGTTTGTTTTGTTTTGTTGAGATGGAGTCTC-ACTCTTTCACCCAGGCTGGAGTGCA 100
             ||| ||     ||||| |    | |||  |||||| |||||  ||| ||||||||||||||||
Qy   659 gtggcgcaattcttggctcactgcaacctctgcctcccggggttcaagtgattattcctgcc 718

Qy   719 cagcctcctgagttagctgggactacaggcgcccaccacagcgctagctaattttttgtatt 778
             ||| |||||||| |||||||| |||||||| ||| ||||| |  || || ||||||| |||
Db   161 CAACCTCCTGAGTAGCTGGGACTACAGGCGCCCGCCACCACGCCCGGCTAATTTTTTGCATT 220

Qy   779 tttagtagaga-tggggtttcaccatggttccgccaggttgatcttctggacctttgtg 837
             |||||||||| | ||||||||||| | ||  |||||| |  |||| ||||  ||| |||
Db   221 TTTAGTAGAGACACGGGTTTCACTGTGTTAGCCAGGATGGTCTCGGATCTCCTGACCCTGTG 280

Qy   838 atctgcctgcctcggcgtcccaagtgtgggattacaggcgtgagccaccacgcccggc 897
             ||| | |||||||||| ||||||| |||||||||||||||||||||||||| |||||
Db   281 ATCCACCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCACGTTCCCGGC 340

Qy   898 --ttattttttaattttgttttgttttgaaatggaatctcactcctgttacccaggctggagt 955
               ||| |||| ||||| ||||| |||| ||| |||| ||||| ||| ||| ||||||||||||
Db   341 TGTTTTGTTTGTTTATTTTTGTTTTGTTTGTTGAGACAGGGTCTCCCTCTGTTGCCAGGCTGGAGT 400

Qy   956 gcaatggccaaatctcggctcactgcaacctctgcctcccggggctcaaagcaattcctg 1015
             ||| ||||  ||||||||||||||||||| | | |||||| ||||||| || |||| | |
Db   401 GCACTGGA-AGGTCTCGGCTCACTGCAACACTGCCCTCCCAGGCTCAAGCGATTCTCGTTG 459

Qy  1016 tctcagcctcccaagcagctgggattacgggcaactgccacacacacccgctaatttttg 1075
             |||| ||||||| ||||||||||||||| ||||  | ||||| || ||||||||| ||||
Db   460 CCTCAGCCTCCCGAGTAGCTGGGATTACAGGCGCGCCACCATGCCCGGCTAAGTTTTG 519

Qy  1076 tatttcattagaggcggggtttcaccacatattgtcaggctggtctcaaactcctgacct 1135
             |||| |||||||||||||||||||||| ||| || |||||||||||||||||||||| |
Db   520 TATTTTTAGTAGAGATGGGGTTTCACCATGTTGGCCAGGCTGGTCTCAAACTCCTGCCTT 579

Qy  1136 caggtgacccacctgcctccaaagtgctgggattacaggcgtgagccaccacctca 1195
             ||| ||| |||||||||| ||  ||| | |||||||||| |||||||| | | |||
Db   580 CAAGTGATCCAACCTGCCTCGGCCTCCCAAAGTGTTGAGATTACAGGAGTGAGCCATTGCG 639

Qy  1196 cccagccggctaatttagataa 1218
             |||  |||||||||| |||| |
Db   640 CCCGGGCAGCAAAAGTTAAACA 662
```

```
FEATURES             Location/Qualifiers
     source          1..970

Qy   599 tattttttattttttaatttttgagacagagtctcaacctcgtcaccaggctggagtgca 658
```

Left alignment

```
RESULT       8

ACCESSION    AQ839814
VERSION      AQ839814.1  GI:6652446
DEFINITION   260L13-C50 CITB Homo sapiens genomic clone 260L13, DNA sequence.
LOCUS        AQ839814             5556 bp    DNA     linear   GSS 31-MAR-2000
```

```
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /clone="IMAGE:5578163"
                    /clone_lib="NIH_MGC_72"
                    /tissue_type="melanotic melanoma"
                    /lab_host="DH10B (phage-resistant)"
                    /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                    Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                    Average insert size 2 kb. Library constructed by Life
                    Technologies."
                    ,G.M.
BASE COUNT    291 a    235 c    240 g    195 t    9 others
ORIGIN
```

```
Query Match            25.6%;  Score 368.8;  DB 10;  Length 970;
Best Local Similarity  78.3%;  Pred. No. 2.2e-71;
Matches    519;  Conservative    0;  Mismatches 133;  Indels  11;  Gaps  6;
```

```
Qy    566 caaagacaatgcgaccactacacctggctaattttttatttttttttgagaca 625
              |||  ||| ||| |||| | ||||  |  ||| |||||||||||||||
Db    904 CAAATCGAGTCTCCCCAAGCACCATTCCGGGGATCAGAATTTTTTTTTTATTTTTTTTTGGTATA 845

Qy    626 gagtctcaactctgtcacccaggctggagtggcgcaaatcttggctcactcgaacc 685
          ||| | |||| ||| |||||||||||| |||||| | | |||||| || |||||
Db    844 GAGTCTCCGG-CTGTCTCCCAGGCTGGAGTGGCAGTGGCGCGATCTCAGTCTCACTGCAAGC 786

Qy    686 tctgcctcccggttcaagttattctcctgcgcccagcctcccctgagtactgggactacag 745
          |||||| || |||||| |||| ||||| |||||| |||| || || | ||| ||||
Db    785 TCCCGCCTTCTGGATTCCCCGCCATTCCTCCTCCGCCTCCNTGAGTAGCTGGGACTACAG 726

Qy    746 gcgcccacca---cgcctagctaattttttttttttattttttagtagtagaga---tgggttcac 799
          || |||| |      |||||| ||||||||||||| |||||| | |||| || |   ||||||
Db    725 GCGCCGGCCACTCGCACCCCGGCTAATTTTTGTATTTTTTTAGTAGAGACGGGGTTTCAC 666

Qy    800 catgttcgccaggttgatcttgatctctggaccttgtgcctccggcctccca 859
          ||||| ||||||||||| |||||| |||| ||||| ||||| || ||||| ||||
Db    665 CGTGTTAGCCAGGATGGTCTCGANCTCCCGACCTCGTGATCCCAGGCCTCAGCCGTGTCA 606

Qy    860 aagtgctgggattacaggcgtgagccaccacgcccgccgctattattttttaattttttgttgtt 919
          || ||||||||||||||| |||| |||||||| |||||| |  ||||| ||||||| |||||
Db    605 AAGTGCTGGGAATTACAGGCATCACCACCACACCCGGCCTTTTTTTT-TTTTTTCTTTTTTTA 547

Qy    920 tgaaatcggaaatccactctgttacccaggctggagtgcaaatggccaaaatctcggctcact 979
          | |||| | ||| |||| ||||||||||||||||||| ||| |||||  | |||||||||||
Db    546 AGACATAGTTTCGCTCT-TGTCGCCCAGGCTGGAGTCAATGGCATGANCATGATCTCGGCTCACC 488

Qy    980 gcaacctctgcctcccgggtcaaggaattctcctgcctcccaagcagactggga 1039
          ||||||||||||||||| | ||| |||||| ||||| || |||||  |||||
Db    487 GCAACCTCCGGCCTCCGGGATTCAAGGCAATTCTCCTGCCTCAGCCTCCCAAGTTGCTGGGA 428

Qy   1040 ttacggcacctgccacacaccccgctaattttttgtattttcattagagaggcgggggtttc 1099
          |||| ||||| || || |  |||||||||||||| |||||||||| | | ||||| ||||||
Db    427 ATACAGGCATGTGCCACCATGCCCACTAATTTTTGTATTTTTAGTAGAGACGGGGTTTC 368

Qy   1100 accatatttgtcaggtgggctctcaaaactcctgacctcaggtgaccacctgcctcagcct 1159
          ||| |  |||||| | ||||| | |||| |||||||| |||| |||| ||| ||||||||
Db    367 ACCACTTTGGCCAGGCTGANCTCGAACTCCTGACCTC--GTGATCCACCCGCCTCGGCCT 310

Qy   1160 tccaaagtgctgggattacaggcgtgagccaccacccagccacctcaccccgggctaatttagataaaa 1219
          ||| ||||||||||||||| | ||| || ||||| |  | ||| ||| | || || |||| |
Db    309 CCCAAAGTGCTGGGATTACAGGTGTGAGCCACCACGCCCAGCCGCGGGGATCAGAGTTTTTA 250

Qy   1220 aaa 1222
          ||
Db    249 AGA 247
```

Right alignment

```
KEYWORDS     GSS.
SOURCE       human.
ORGANISM     Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 5556)
AUTHORS      Carpten,J.D., Makalowska,I., Robbins,C.M., Scott,N., Sood,R.,
             Connors,T.D., Bonner,T.I., Smith,J.R., Faruque,M.U., Stephan,D.A.,
             Pinkett,H., Morgenbesser,S.D., Su,K., Graham,C., Gregory,S.G.,
             Williams,H., McDonald,L., Baxevanis,A.D., Klingler,K.W. and Landes
             ,G.M.
TITLE        A 6-Mb high-resolution physical and transcription map encompassing
             the hereditary prostate cancer 1 (HPC1) region
JOURNAL      Genomics 64 (1), 1-14 (2000)
MEDLINE      20175426
COMMENT      Contact: Carpten JD
             Cancer Genetics Branch
             National Human Genome Research Institute/National Institutes of
             Health
             Bldg. 36, Room 3004, 36 Convent Drive, Bethesda, MD
             Tel: 301 435 5626
             Fax: 301 435 5465
             Email: jdc@nhgri.nih.gov
             Class: shotgun.
FEATURES             Location/Qualifiers
     source          1..5556
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="260L13"
                     /clone="CITB"
BASE COUNT    1500 a    1016 c    1002 g    2035 t    3 others
ORIGIN
```

```
Query Match            25.5%;  Score 367.6;  DB 12;  Length 5556;
Best Local Similarity  80.2%;  Pred. No. 3.3e-71;
Matches    483;  Conservative    0;  Mismatches 109;  Indels  10;  Gaps  4;
```

```
Qy    604 ttatttttaattttttgagacagagtctcaactctgtcacccaggctgggagtgcagtggc 663
          |||| ||| ||||||| ||||| |||||||| || | |||||||||||||||| | ||||
Db   3145 TTAACTTTTTTTTTTTTTTTGAGACAGAGTCTCGCTCTGTCGCCCAGGCTGGAGTGCAGTGGC 3204

Qy    664 gcaaatcttggctcactgcaacctctgcctcccgggttcaagtgattattctcctgcctcagcagcc 723
          ||| |||| ||||||||||||| |||||| | ||| | ||||| ||| || |||||||||||||
Db   3205 GCGATCTCGGCTCACTGCAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCC 3264

Qy    724 tcctgagtagctgggactacaggcgc---ccaccacgcccagcctagctaatttttttgtattt 780
          ||| |||||||||||||||||||| |   ||||||  | || ||| ||||||||||||||||
Db   3265 TCCCGAGTAGCTGGGACTACAGGCGTCATGCCACCACGCCCAGCTAATTTTTTGTATTTTTA 3324

Qy    781 tagtagagatgggtttcaccatgttgccaggttgatcttggacctgtgatc 840
          ||||||||| ||||||||||| || |||||| |||||||||||| ||||||
Db   3325 GTAGGGACA-GGGTTTCACCATGTTGGCCAGGATGGTCTCGATCTCTTGACCTCGTGATC 3383

Qy    841 tgcctgcctcggcctcccaaaagtgctgggattacaggccacacgcccggctta 900
          |||||||||| ||||||| ||||||||||||||||||| |||| |||||  ||||
Db   3384 CGCCTGCCTCGGCCTCCCAAAAGTGCTGGGATTATAGGCGTGAGCCACTGTGCCCGGCCTC 3443

Qy    901 tttttaattttgtttgtttgaaatggaatctcactctgttacccaggctgggagttcaat 960
          ||||| ||||| |||||||  ||||| ||||||| ||||||||||||||  ||| || |
Db   3444 TTTTT--TTTTTTTTTTTTGAGATGGGAGTCTCGTGTCTCCCAGGCTGGAGTGCAGT 3501

Qy    961 ggcaaatctcggctcactgcaagctctgcctcccgggttcaagcgattctcctgtctcca 1020
          |||| |||| ||||||||||| | |||||| | ||| | ||||| |||| || |||||
Db   3502 GGCACACAATCTTGGCTCGACTGCAACCTTTGCCTCCCGAGTTCAAGCGATTCTCCTGCCTCA 3561

Qy   1021 gcctcccaagcagctgggattacgggcacctggccacaccccgctaattttttgtatttatt 1080
          |||| | ||||| |||| ||| |||||||| | | |||||| |||||||| || |
Db   3562 GCCTCCCAAGTAGCTGGGATTA---CACTCACCACCAACACCCGGCTAATTTTTGTATTT 3617

Qy   1081 tcattagaggcggggtttcaccaccatatttgtcaggctggtctcaaactcctgacctcaggt 1140
          |||| ||||| |||||||||| | |||  || | |||| |||||||||||||||||||| |||
Db   3618 TTAGTAGAAATGGGGTTTCACCATGTTGGCCAGGCTGGTCTCGAACTCCTGACCTCAAGT 3677
```

```
Qy    1141 gacccacctgcctcaagccttcaaaagtgctggggattacacggcgtgagccaacggcccggc 1200
           ||||||||||||||||||||||||||||||||||||||||||||| ||||| ||| ||||| |
Db    3678 GATCCAACCTGCCTCAGCCTCCCAAAAATGTTGGGATTACAGGCGTGAGCGCACCACCCCGG 3737

Qy    1201 cc 1202
           ||
Db    3738 CC 3739
```

Query Match   25.4%; Score 366.6; DB 9; Length 839;
Best Local Similarity   81.4%; Pred. No. 7e-71;
Matches 499; Conservative   0; Mismatches 101; Indels 13; Gaps 6;

```
Qy    602 ttttatttttaattttttggacaagaagtctcaactctgtcaccaggctgggagtgcaagtg 661
          |||| ||||| || |||| |  ||||||||  |  ||||||||||||||||||  |||||| |
Db     66 TTCTTTTTTTTAATTTTTTGAGCAGGAGTCT-GGCTCTGTCACCAGGCTGGAGTGCAGCG 124

Qy    662 gcgcaatcttggctcactgcaacctctgcctcccgggttcaagttattctctcctgcgccag 721
          ||| | |||||||||||||||  ||||| |  |||||||||||  || |||| |||| || |
Db    125 GCGCGATCTCGGCTCACTGCAAGCTCCTGCCTCCCGGGTTCACCCATTCTCCTGCCTCAA 184

Qy    722 cctcctgagtagctgggactacaggcgc---cccacacgcctagctaattttttgtatt 778
          |||||||||||||||||||||||||||     | ||| |||| ||||| |||||  |||
Db    185 CCTCCCGAGTAGCTGGGATTACAGGCGTTCGCCACCACGCCCGGCTAATTTTTTGTATT 244

Qy    779 tttagttagagatggggg-ttcaccatgttcgcgccaggttgatctctgtgacctgtg 837
          |||||||||||||||||  |||||||||||||||||||||||||| |||||||| |||
Db    245 TTTAGTAGAGACGGGGTTTCACCGTGTTAGCCAGGATGGGGTCTCGGATCTTCTGACCTCGTG 304
```

Query Match   25.0%; Score 360; DB 9; Length 1148;
Best Local Similarity   80.9%; Pred. No. 1.9e-69;
Matches 491; Conservative   0; Mismatches 99; Indels 17; Gaps 6;

```
Qy    838 atctgcctgcctcggcctcccaaagtgctgggattacacggcgtgagccaacggcccggc 897
          ||||||||||||||||||||||||||| ||||||| ||||||||||||||| ||||| |
Db    305 ATCTGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGACCACCGCACCACCCGGGC 364

Qy    898 ttattttaatttttgtttgaaatggaatctcactcttgttaccaggctgggagtgc 957
          ||||| || |||| ||||| |||  |||||||||| ||| ||||||||||||  ||||
Db    365 CTCTTATTTTTTTTT------TTTGAGATGGAGTCTCACACTGTCACCAGGCTGGAGTGC 418

Qy    958 aatggccaaatctctcggctcactgcaacctctgccctcccgggttcaagcggattctctgtc 1017
          ||| | |||||||||  |||||| | |||| ||||| | ||||||| |||| ||||  ||| |
Db    419 AGTGGAGCGGATCTCGGCTCACTGCAACCTCCGCCCGCTCGGGTTCAAGCGATTCTCCTGCC 478

Qy    1018 tcagcc-tcccaagcagctgggattacagggccacctgccacaccacccgccgcta-attttg 1075
           ||||| |||||||||||||||||||||||| |||||||||||||||||||||| |||| |||||
Db    479 TCAGCCTTTCCAAGTAGCTGGGATTACAGGTGCCCACCACCACGCCCGGCTAGTTTTTTG 538

Qy    1076 tattttccattagagcgggggtttcaccatattttgtcaggctggtctcaaactcctgacct 1135
           |||||||||||||| |||  ||| |||||| | |  |||  ||||||||| |||||||||||
Db    539 TATTTTTAGGCAAGATGGGGTTTTACCATGTCTGGCCAGGCTGGTCTCGAACTCCTGACCT 598

Qy    1136 caggtgaccccaacctgcctcagccctcccaaagtgctgggattacaggcgtgagccaccca 1195
           |||||||||||||||||||||||||||||||| |||||||| ||||||||| ||||||||||
Db    599 CAGGTGAACCGCCGCCCACCTTAGCNCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCATA 658

Qy    1196 cccagccggctaa 1208
           |||||||| |||
Db    659 CCTGGGCCGCAAAA 671
```

```
RESULT 11
AG150197/c
LOCUS       AG150197              667 bp    DNA      linear   GSS 09-JAN-2002
DEFINITION  Pan troglodytes DNA, clone: RP43-012023.T7, genomic survey
            sequence.
ACCESSION   AG150197
VERSION     AG150197.1  GI:16679875
KEYWORDS    GSS; GSS (genome survey sequence).
SOURCE      Pan troglodytes male lymphocytes'DNA, clone_lib:RPCI-43 Chimpanzee
            Male BAC Library clone:RP43-012023.T7.
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1  (sites)
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     BAC end sequences of library RPCI-43
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 667)
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe.H. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:chimpes@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     Clones are derived from the chimpanzee BAC library RPCI-43 This BAC
            end was generated during the R&D process and may have higher chance
            of clone tracking errors.
            PRIMERS
            Sequencing: T7
FEATURES             Location/Qualifiers
     source          1..667
                     /organism="Pan troglodytes"
                     /db_xref="taxon:9598"
                     /clone="RP43-012023.T7"
                     /sex="male"
                     /cell_type="lymphocytes"
                     /clone_lib="RPCI-43 Chimpanzee Male BAC Library"
BASE COUNT      192 a    163 c    184 g    128 t
ORIGIN

Query Match        25.0%;   Score 359.8;   DB 12;   Length 667;
Best Local Similarity   81.3%;   Pred. No. 2.3e-69;
Matches  495;   Conservative   0;   Mismatches  97;   Indels  17;   Gaps  6;
```

```
Qy   599 tattttatttttaatttttgagacagagtctcaactctgtcacccaggctggagtgca 658
             ||||||||| ||||||||||||||||||||||| ||||||||||||||||||||| ||
Db   608 TTTTTNTNTTTTTTTTTTTTTTNTTTGAGACGGCAGTCTC-ACTCTGTCGCCCAGGCTGCA 550

Qy   659 gtgggcaatcttggctcactgcaacctcccggttcaagttattctcctggcc 718
             ||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db   549 GTGGCGTGANTCACAGCTCACTGCAACCTCTGCTCTNNGGGTTCAAGCAAATTCTTCTCGCC 490

Qy   719 -cagcctctgagtagtagggaccacaggggcccaccacgcctagctaat-tttttgta 776
             |||||||||||||||| ||||||||||||||||||||| |||||||| ||||||| ||
Db   489 TCAGCCTCTGAAGTAGCTGGGACTACAGGCGCNCCGCACNNAGNACGGANTCTTCTGCGG 430

Qy   777 tttttagtagaga-tggggttcaccatgttcgccaggttgatctctggaacttg 835
             ||||||||||| | ||||||||||||| |||||||||||| || |||||||||||
Db   429 TTTTTAGTAGAGACAGGGGTTTCATCANTGTTGGCCAGGCTGTCTCTNGAACTCCCTG 370

Qy   836 tgatctgctgcttcggccctcccaaagtgctggggattacaggcgtgagccacacgcccg 895
             ||||||||||||| || ||||||||||||||||||||| |||||||||||||||||||||
Db   369 TGACCACCACCACGGACTNCCTCAAGTGCTGGGGATTACAGGCGTGAGCCACCACCACCCG 310

Qy   896 gc---------ttattttttttgttttgaatctcactctgttacc 944
               |          ||||||| |||||||||||||||||||| |||
Db   309 GCCTCATTTGCCAATTCTTTTTTTTTTTCTTTGAGATGGAGTCTCACTCTNTATTGCC 250

Qy   945 caggctggagtgcaaatgcgcaaactctcggctcactgcaacctctgcctcccggttcaag 1004
             || |||||||||| ||||| || ||||||||||||||||||||||||||||||||||||||
Db   249 CAGGCTGGAGTGCTGTGGCGCAAATCTCGGCTCACTGCAACCTCTGCCTCCCGGGATTCAAG 190

Qy  1005 cgattcctcgttctcagcctcccaagcagcgtgggattacgggcacctgccacacaccccc 1064
             ||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db   189 TGATTCTCTGCCTGCCTTCCGAGTAGCTGGGATTACAGGCATGCCACCACCATGCCCA 130

Qy  1065 gctaatttttgtattttcattagaggcggggtttcaccatatttgtcaggctgggtctcaa 1124
             ||||||||||||||||||||||||||||||||||||||||||||| ||||||||| |||||
Db   129 GCTAATTTTTGTATTTTTACTAGAGACAGGGTTTCACCANCATTGGTCAGGCTGGTCTCGAA 70

Qy  1125 actcctgacctcaggtgacccacctgcctcagccttcccaaagtgctgggattacaggcgt 1184
             |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db    69 ACTCCTGACCTC--GTGATCCGCCCGCCTCCCAAAGTGCTGGGATTACAGGCGT 12

Qy  1185 gagccac 1191
             |||||||
Db    11 GAACCAC 5
```

```
RESULT 12
AG186062/c
LOCUS       AG186062              687 bp    DNA      linear   GSS 09-JAN-2002

Query Match        25.0%;   Score 359.8;   DB 12;   Length 667;
Best Local Similarity   81.3%;   Pred. No. 2.3e-69;
Matches  495;   Conservative   0;   Mismatches  97;   Indels  17;   Gaps  6;

Qy   601 tttttatttttaatttttgagacagagtctcaactctgtcacccaggctggagtg--- 656
             |||||||||| ||||||||||||||||||||| |||||||||||||||||||||
Db   659 TTTTTTTTTTTTTTTTTTTTTAAGACAGAGTTTCGTCTCTTGTCCTCTTGTTGCCCAGGCTGGA 601

Qy   657 cagtggcgcaatcttggctcactgcaacctctgcctcccgggttcaagttattctcctgc 716
             |||||||| || ||| |||||||||||||||||||||||||||||||| |||||||||||
Db   600 CAGTGGGATCGATGTCGGCGGCTCACTGCCAGCTCCACCTCCCAGGTTCACGCGCCATTCTCCTGC 541

Qy   717 cccagcctcctgagtagctgggactacaggcgcgccacacacgcctagc--taatttttg 774
             |||||||||||||| |||||||||||| ||||| |||| ||||| || |  |||||||| |
Db   540 CTCGGCCTCCCAAGTAGCTGGGATTACAGGCGCCCACCACCACGGAGATGGTTCGATCTCGGACT 481

Qy   775 tattttagtagagatgggg--ttcaccatgttcgccaggttgatctctggacct 833
             |||||||||||||||| ||  ||||||||||||| || |||||| | |||||||
Db   480 TATTTTTAGTAGACAGGGGTTTCACTATGTTAGCCAGGATGGTCTCGATCTCCTGACCT 421

Qy   834 tgtgatctgctgcctcggccctcccaaagtgctggggttacagggccacacgcc 893
             |||||||||||| || ||||||||||||||||||||||||||||||| |||||||
Db   420 TGTGATCCACCCGTCTCGGCCTCCCAAAGTGCTGGGATTACAGGCATGAGCCACCGTGCC 361

Qy   894 cggcttatttttgttttgaaatcactctgttacccaggctgga 953
             ||| ||||||||||||||||| |||||||||||||||||||||
Db   360 TGGCCTT-----TTTTTTTTTTTAAGCACAGAGTTTTGCTCTTGTTGCCCAGGCTGGA 308

Qy   954 gtgcaatgcgcaaatctcggctcactgcaacctctgcctcccgggctcaagcgattctcc 1013
             ||||||| || |||||||||||||||||||||||||||||||||||||||| |||||||||
Db   307 GTGCAATGCGCGATCTCAGCTCACTGCAACCTCCGCCTCCCGGGTTCAAGGGATTGTCC 248

Qy  1014 tgtctcagcctcccaagcagcgtgggattacgggcacctgccacacaccccgctaattt 1073
             |||||||||||||||||||||| |||||||||||||| |||||||||||||||||||||
Db   247 TGCCTCAGCCTCCTGAGTAGCTGGGATTACAGGCATGCCACCACCCCCGCTAATTTT 188

Qy  1074 tgtattttcattagaggcgggggtttcaccacatattgtcaggctggtctcaaactcctgac 1133
             |||||||||||||||||| ||||||||||||| |||||||||| ||||| |||||| |||||
Db   187 TGTATTTTTAGTAGAGACGGGGTTTCACCATGTTGGTCAGGCTGGTCTCGGAACTCCTAAC 128

Qy  1134 ctcaggtgacccatctgcctcagcctcccaaagtgggattacaggcgtgagccacct 1193
             ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   127 CTC--GTGATCCGCCTGCCTCACCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCA 70

Qy  1194 cacccagcc 1202
             |||||||||
Db    69 TGCCCAGCC 61
```

DEFINITION Pan troglodytes DNA, clone: RP43-060F11.T7, genomic survey
            sequence.
ACCESSION   AG186062
VERSION   AG186062.1 GI:16715742
KEYWORDS   GSS; GSS (genome survey sequence).
SOURCE   Pan troglodytes
            Pan troglodytes male lymphocytes DNA, clone_lib:RPCI-43 Chimpanzee
            Male BAC Library clone:RP43-060F11.T7.
ORGANISM   Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1 (sites)
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE   BAC end sequences of Library RPCI-43
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 687)
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE   Direct Submission
  JOURNAL   Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:chimpbes@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT   Clones are derived from the chimpanzee BAC library RPCI-43 This BAC
            end was generated during the R&D process and may have higher chance
            of clone tracking errors.
            PRIMERS
              Sequencing: T7
FEATURES   Location/Qualifiers
     source   1..687
                     /organism="Pan troglodytes"
                     /db_xref="taxon:9598"
                     /clone="RP43-060F11.T7"
                     /sex="male"
                     /cell_type="lymphocytes"
                     /clone_lib="RPCI-43 Chimpanzee Male BAC Library"
                     LIBRARY
                       Vector   : pBACe3.6
                       R.Site 1 : EcoRI
                       R.Site 2 : EcoRI
BASE COUNT   207 a   161 c   182 g   137 t
ORIGIN

Query Match   24.98%; Score 358.8; DB 12; Length 687;
Best Local Similarity   81.2%; Pred. No. 3.8e-69;
Matches   492; Conservative   0; Mismatches 102; Indels   12; Gaps   6;

Qy   601 tttttattttttaattttttgagacagagtctcaactctgtcacccaggctggagtgcagt 660
              |||||||||| ||||||||| ||||||||||||| |||||||||||||||||||| ||||||
Db   655 TCTTTTGTTTTTTTTTTTTAGACAGAGTCT-TGCTCTGTCACCCAGGTTGAAGTCAAGT 597

Qy   661 gggcgcaatcttggctcactgcaacctctgcctcccgggttcaagtgattattctctgccca 720
              ||||| ||||| ||||||||||| |||||| |||| || ||| ||||| ||||||||||
Db   596 GGGCGGGATCTCGGCTCACTGCGAAGCTCCACGTCACCAGGTTCACGCCATTCTCGTCGCTCG 537

Qy   721 gcctcctgagttagctgggactacaggcgcccaccacgcccaggctagc--taattttttgtatt 778
              ||||||| || ||| ||||| ||||||||| || ||||| |||||||| |||||||||||
Db   536 GCCTCGCAAGTAGCTGGGATTACAGGCCCCCACCACCACGGCCGGCTAATTTTTGTATTTTT 477

Qy   779 tttagtagagatggggg-ttcaccatgttccgccaggttgatctctgaccttgtg 837
              ||||||| ||||||| ||||||||| || |||||||||| |||||||||||||
Db   476 TTTAGTAGAGACGGGGTTTCACCATGTTAGCCAGGATGGTCTCGATCTCCTGACCTTGTG 417

Qy   838 atctgcctcgcctcggcctcccaaagtgctgtggattacaggcgtgagccacacgcccggc 897
              | ||||||||| ||| ||||||||||||||||||||||||||| ||||||||| | |||||
Db   416 ATCCACCTGTCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCGCCCTGGC 357

Qy   898 ttatttttttttttgtttgttttgaaatggaatctcactctgttacccaggctggagtgc 957
              |||||| |||| |||||||||| | ||||||||||||||||| |||||||||||||||||
Db   356 TTTTTTTTTTTTTTATTTAAGACAGAGTTTCGCTCT-----TGTTGCCCAGGCTGGAGTGC 302

RESULT 13
LOCUS   BF346320   679 bp   mRNA   linear   EST 22-NOV-2000
BF346320/c
DEFINITION 602018525F1 NCI_CGAP_Brn67 Homo sapiens cDNA clone IMAGE:4153999
            5', mRNA sequence.
ACCESSION   BF346320
VERSION   BF346320.1 GI:11293915
KEYWORDS   EST.
SOURCE   human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 679)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: David N. Louis, M.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9423 row: a column: 08
            High quality sequence stop: 661.
FEATURES   Location/Qualifiers
     source   1..679
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4153999"
                     /clone_lib="NCI_CGAP_Brn67"
                     /tissue_type="anaplastic oligodendroglioma with 1p/19q
                     loss"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally. Primer: oligo dT.
                     Average insert size 2.3 kb. Constructed by Life
                     Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT   200 a   161 c   191 g   127 t
ORIGIN

Query Match   24.8%; Score 357.8; DB 10; Length 679;
Best Local Similarity   79.9%; Pred. No. 6.3e-69;
Matches   485; Conservative   0; Mismatches 112; Indels   10; Gaps   5;

Qy   600 atttttattttttgagacagagtctcaactctgtcacccaggctggagtgcag 659
              ||| ||| ||||||||||||||||||||||||||||| |||||||||||||||
Db   676 ATTCTTTGTTTTTTCGCTATTGAGACGAGAGTTTTGCTCTGTTGCCCAGGCTGGAGTCAG 617

RESULT 14

```
Qy    660 tggcgcaaatcttggctcaactcgctcccgggttcaaagttattctcctgccc 719
          |||  ||||||| |||||||||||||||||||||||||||||||||||||||||||
Db    616 TGGCATGATCTTGGATCACTGAACTCTCCCTGGTTCAAGCAGTCTCTGCCTC 557

Qy    720 agcctccgagtagctgggactaacaggcgccacacagccagctaat--ttttgtatt 778
          ||||||| ||  |||||||||| ||||||| ||||||||||||| ||| |||||||||
Db    556 AGCCTCCGAGTAGCTGGGATTACAGGCACATCACCACCACCACCCTG TAATTTTGTATT 497

Qy    779 tttagtagagatgggg-ttcaccatgttcgccaagttgatcttctggacct--tg 835
          ||||||||||||||||  |||||||||||||||||||||||||||||||| ||
Db    496 TTTAGTTAGAGACGGGGTTTCGCCATGTTCACCAGGCTGGTCTTGAACTCCTGAACCTCAGG 437

Qy    836 tgatctgcctcggcctcccaaaagtgctgggattacaggcgtgagccaccacgccccg 895
          |||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db    436 TGATCCACCCACCTCAGCCTCCCGAAGTGCTGGGATTACAGGCATGAGCCACTGTGCCCT 377

Qy    896 gcttattttaattttgttgtttgaaatggaatctcacctcgttacccaggctggagt 955
          |||||||||||||||||||||||||     |||||||||||||||||||||||| ||
Db    376 GATTATTTTTTATTTTT-----TTTGAGACAGAGTCTTGCTCTGTCGCCCAGGCTGGAAT 322

Qy    956 gcaatggcgcaaatctcggctcactgcaacctctcgcctcccgggttcaagcgattcctg 1015
          ||||| ||||||||||||||||||||||||||||||||||||||| |||||||||||
Db    321 GCAGTGGCACGATCTTTGGCTCACTGCAACCTCTGCCTCCCGGGTTTAAGCGATTCTCATG 262

Qy    1016 tcctcagctctcccaagcagctgggattacgggcacctgccaccacacccggctaattttg 1075
           ||||||||||||||||||||||||||||  ||||||||||||||||||||| ||||||||
Db    261 CCTCAGCCTCCCAAGTAGCTGGGATTACAGGCGCCCACCACCACGCCCGGCTAATTTTTG 202

Qy    1076 tattttcattagagggcggggtttcaccatatttgtcaggctggtctccaaactcctgacct 1135
           ||||||||||||||||||||| |||||||||| ||||||||||||||  |||||||||||
Db    201 TATTTTAGTAGAGANGAGGTTTCACCATGTTGGCCAGGCTGGTCTCAAACTCCTCACCT 142

Qy    1136 caggtgacccacctgccctcagccttccaaagtgctgggattacaggcgtgagccacctca 1195
           |||||| ||||||||||||||||| |||||||||||||||||||||||||||||||||
Db    141 CAGGTGATCCGCCCGCCTCGGCCTCCCAAAGT-CTGGGATTACAGGCGTGAGCCACCATCGTG 83

Qy    1196 cccagcc 1202
           ||| |||
Db    82 CCTCGGCC 76
```

RESULT 14
LOCUS   AQ583981   722 bp   DNA   linear   GSS 07-JUN-1999
DEFINITION   RPCI-11-452N24.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-452N24
              , DNA sequence.
ACCESSION   AQ583981
VERSION   AQ583981.1   GI:5011091
KEYWORDS   GSS.
SOURCE   human.
ORGANISM   Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 722)
AUTHORS   Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and Venter
              ,J.C.
TITLE   Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
              Map Building
JOURNAL   Unpublished (1997)
COMMENT   Other_GSS: RPCI-11-452N24.TV
              Contact: Shaying Zhao, William Nierman, Mark Adams
              Department of Eukaryotic Genomics
              The Institute for Genomic Research
              9712 Medical Center Dr., Rockville, MD 20850
              Tel: 301 838 0200
              Fax: 301 838 0208
              Email: hbe@tigr.org
              Clones are derived from the human BAC library RPCI-11. For BAC
              library availability, please contact Pieter de Jong
              (pieter@dejong.med.buffalo.edu). Clones may be purchased from
              BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
              Research Genet cs (info@resgen.com). BAC end search page:
              http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.

FEATURES   Location/Qualifiers
     source   1..722
                /organism="Homo sapiens"
                /db_xref="GDB:7673519"
                /db_xref="taxon:9606"
                /clone="RPCI-11-452N24"
                /clone_lib="RPCI-11"
                /sex="Male"
                /cell_type="Lymphocytes"
                /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                RPCI11 Human Male BAC Library"

Seq primer: SP6
Class: BAC ends.

BASE COUNT   220 a   168 c   199 g   135 t
ORIGIN

Query Match   24.8%;   Score 357.4;   DB 12;   Length 722;
Best Local Similarity   80.7%;   Pred. No. 7.7e-69;
Matches 506; Conservative   0; Mismatches 106;   Indels 15;   Gaps 7;

```
Qy    587 ctggctaattttttattttttattttttaattttttgagacagagtctcaactctgtcacca 646
          ||||||||| |  ||||||||||||||||||||||||||||||||||||||||||||||
Db    653 CTGGCTTTTCCAATTATTATTATTTATTTTTTTTTTTTTGAGTTGGAGTCT--GGCTCTGTCGCCCAG 594

Qy    647 ggctggagtgcagtggcgcaaatcttggctcactgcaacctctgcctcccgggttcaagtt 706
          ||||||||||||||| ||||||||||||||||||||||| ||||||||||||||||||
Db    594 GGCTGGAGTGCAGTGGCGTGATCTCGGCTCACTGCAAGCTCTGCCTCCCGGGTACAGCGCC 535

Qy    707 attctcctgcccagcctcctgagtagctgggactacaggcgcccaccac--gcctagc 763
          |||| | |||||||||||||||||||| |||||||||||| ||||||||  ||||||
Db    534 ANTCTCCTGCCTCAGCCTCCCAAGTAGCAGGGACTACAGGCGCCCACCACCATGCCTGGC 475

Qy    764 taatt---tttttgtatttttagtagagatgggg-ttcaccatgttcgccaggttgatct 819
          ||||||    ||||||||||||||||||||||||  ||||||||||||||||||||||||
Db    474 TAATTTTTTGTTTTTGTATTTTTAGTAGAGACGGGGTTTCACCGTGTTAGCCAGGATGAGTCT 415

Qy    820 tgatctcctgaccttgtgatctgcccgccttattt----taattttgtttgagatggaatctcac 879
          |||||||||||| ||||||||||||||||||||       |||||||||||||||||||
Db    414 CGATCTCCTGACCTCGTGATCTGCCCGCCTTGGCCTCCCAAAGTGTTGGGATTACAGGATG 355

Qy    880 tgagccacagcccccggcgcttatttt----taatttttgttttgaaatggaatctcac 935
          |||||||||||||| ||||||||||||    |||||||||||| ||||||||||||||||
Db    354 TGAGCCACTGTGCCCAGGCTGGCTTTCCAATTCTTTTCTTTTTTTTTGAGATGGAGTCTCAC 295

Qy    936 tctgttacccaggagtgcagtggcgcaatggcgcaaatctcggctcacttgcaacctctgcctccc 995
          ||||||||||||| |||||| ||  |||| |||||||||||||||||| ||||||||||||||
Db    294 TCTG-TCGCCCAGGCTGGAGTGCAGTGGCGCAATCTCAGCTTAGTGCAACCTCTGCCTCC 236

Qy    996 gggctcaagcgattccctcgtctccagccctcccaagcaagctgggattacgggcaactgcca 1055
          |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db    235 GGGTTCAAGTGATTCTCCTGCCTCAGCCTCCCGAATAGCTGGGATTACAGGCACCCGGCCA 176

Qy    1056 ccacacccgctaattttgtatttcattagagcggggggtttcaccatatttgtcaggc 1115
           |||||| |||||||||||| ||||||||||||||||||||||||||||||||||||||
Db    175 CCACGCGTTGGCTAATTTTTGTATTTTTAGTAGAGACGGGGTTTCACCATGTTGGCCAGGA 116

Qy    1116 tggtctccaaactcgacctcaggtgacccacctgccctcagccttccaaaagtgctgggat 1175
           |||||||| |||||||||||||||||||||||||||||||||||||||||||||| ||||
Db    115 TGGTCTTGATCTCCTGACCTCA--TGATCCGCCCGCCTCGGCCTCCCAAAGTGATGGGAT 58

Qy    1176 tacaggcgtgagccacctcaccagcc 1202
           |||||||||||||||||||| || ||
Db    57 TACAGGCCATGAGCCACGGGCGCCAACC 31
```

RESULT 15
LOCUS   BC009590   1259 bp   mRNA   linear   HTC 22-OCT-2001
DEFINITION   Homo sapiens, Similar to neuronal thread protein, clone
              IMAGE:4102657, mRNA.
ACCESSION   BC009590
VERSION   BC009590.1   GI:16307027

```
KEYWORDS     HTC.
SOURCE       human.
ORGANISM     Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 1259)
AUTHORS      Strausberg,R.
TITLE        Direct Submission
JOURNAL      Submitted (29-JUN-2001) National Institutes of Health, Mammalian
             Gene Collection (MGC), Cancer Genomics Office, National Cancer
             Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
             USA
REMARK       NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT      Contact: MGC help desk
             Email: cgapbs-r@mail.nih.gov
             Tissue Procurement: ATCC
             cDNA Library Preparation: CLONTECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Sequencing Group at the Stanford Human Genome
             Center, Stanford University School of Medicine, Stanford, CA 94305
             Web site:       http://www-shgc.stanford.edu
             Contact:  (Dickson, Mark)  mcd@paxil1.stanford.edu
             Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
             R. M.
             Clone distribution: MGC clone distribution information can be found
             through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
             Series: IRAL Plate: 21 Row: n Column: 9
             This clone was selected for full length sequencing because it
             passed the following selection criteria: Similarity but not
             identity to protein
             This clone has the following problem: frame shifted.
FEATURES             Location/Qualifiers
     source          1..1259
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4102657"
                     /tissue_type="Brain, glioblastoma"
                     /clone_lib="NIH_MGC_57"
                     /lab_host="DH10B"
                     /note="Vector: pDNR-LIB"
BASE COUNT      383 a    230 c    303 g    343 t
ORIGIN
```

```
·    Query Match          24.7%;  Score 356.6;  DB 11;  Length 1259;
     Best Local Similarity   80.6%;  Pred. No. 1.1e-68;
     Matches  491;  Conservative  0;  Mismatches 109;  Indels  9;  Gaps  6;

Qy     602 ttttattttttaatttttgagacagagtctcaacctctgtcaccaggctggagtgcagtg 661
           ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db     890 TTTTTTTTTTTTTTTTTTTTGAGACAGAGTTTTCGTTGTTGCCCAGGCTGGAGTGCAAtG 831

Qy     662 gcgcaatcttggctcactgcaacctctgcctcccgggttcaagtgattattctcctgcccag 721
           | |||||||||||| | ||| ||||||||||||||| ||||||||||||||||||||||||
Db     830 GTGCAAtCTCGACTCACTGCAACCTCACCTCCTCGGGTTCAAGCTAATTCTCCTGCCTCAG 771

Qy     722 cct-cctgagtagctgggactacaggcgc---ccaccacgcctagctaattttttgtat 777
           ||| |||||||||||||||||||||||||    |||||||||||||||||||||||||||
Db     770 CATCCCAAAGTRAGCTGAGATTACAGGCACCCGCGCCCACCACGCCCAGCTAATTTTTTGTAT 711

Qy     778 ttttagttagagatggggg-ttcaccatgttccgccaggttgatcttgactctggac--ctt 834
           |||| | |||| ||||||| |||||||||||||||||||||||||||||||||||||   |||
Db     710 ATTCTGTTGAGGTGGGGTTTCACCATGTTGGCCAGGCTGGTCTTGAGCTCCTGACTTCAG 651

Qy     835 gtgaatctgcctcggcctcccaaaagtgtggggattacaggcgtgagccacacgccc 894
           ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db     650 GTGATCCACCCACCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCGCC 591

Qy     895 ggcttatttttgttctgttttgaaatggaatccacctctgttacccaggctggag 954
           ||||||     |||||||||||||||||||||||||||||||||||||||||||||
Db     590 GGCCAtTCCCACAtAATTTTCTTTTTGTTGACAGAGTCTTGCTCTGACACCCAGGCTGGAC 531

Qy     955 tgcaatgccaaatctcgggctcactgcaacctctgcctcccgggctcaag-cgattctcc 1013
            ||| ||| |||||||||||||||||||||||||||||||||||||||||| |||||||||
Db     530 TACAGTGGTGCGAtCTCAGCTCACTCACTGCAACCCTCTGCCTCCCCTGGGTTCAAGCCGAtTCTCC 471

Qy    1014 tgtctcagcgctcccaagcagctgggattacgtggcacctgccacacacacccgctaattt 1073
           || ||||||||||||||||||||| ||||||||||||||||||||||||||||||| |||||
Db     470 TGCCTCAGCCTCCCGAGTAGCTGGGAtTACAGGCAtGGCCACCACGCCCTGGCTAAtTTTT 411

Qy    1074 tgtattttccattagagcgggttcaccatatttgtcaggctggtctcaaactcctgac 1133
           |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db     410 GGTATTTTTAGTAGAGAtGGGGTTTCACCATGTTGGTCAGGCTGGTCTCAAACTCCTGAC 351

Qy    1134 ctcaggtgacccacctgcctcagcctcccaaagtgctgggattacaggcgtgagccacct 1193
           ||||||| ||||| || ||||||||||||||||||||| |||||||||||||||||||||
Db     350 CTCAGGT-ATCTACCCACCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACTG 292

Qy    1194 cacccagcc 1202
           |||| ||||
Db     291 CGCCCGACC 283
```

```
Search completed: April 24, 2002, 19:05:57
Job time: 7362 sec
```

OM nucleic - nucleic search, using sw model

Run on: April 24, 2002, 18:39:56 ; Search time 207.25 Seconds
                 (without alignments)
                 11945.914 Million cell updates/sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Title: US-09-964-678A-1
Perfect score: 1442
Sequence: 1 ttttttttttgagatggag..........ttaaacaaagcttagagca 1442
Scoring table: IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0
Searched: 1736436 seqs, 858457221 residues
Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 listing first 45 summaries

Database :   N_Geneseq_032802:*
1: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query % Match	Length	DB	ID	Description
1	1223.4	84.8	1418	17	AAT27738	Neural thread prot
2	1078.6	74.8	1381	15	AAQ77883	Neural thread prot
3	1077	74.7	1381	17	AAT27765	AD 10-7 human neur
4	517.2	35.9	3374	22	AAK69279	Human immune/haema
5	517.2	35.9	3747	22	AAK73093	Human immune/haema
6	511.6	35.5	5262	22	AAK71768	Human immune/haema
7	511.6	35.5	5262	22	AAK71769	Human immune/haena
8	489.6	34.0	24167	22	ABA16132	Human nervous syst
9	478.2	33.2	31853	22	AAI98993	Human excretory re

RESULT   1
AAT27738

ID   AAT27738 standard; DNA; 1418 BP.
XX
AC   AAT27738;
XX
DT   13-NOV-1996 (first entry)
XX
DE   Neural thread protein coding sequence.
XX
KW   Neural thread protein; NTP; diagnosis; detection;
KW   Alzheimer's disease; neuroectodermal tumour; malignant astrocytoma;
KW   monoclonal antibody; binding fragment; ds.
XX
OS   Homo sapiens.
XX
FH
FT   Key   Location/Qualifiers
FT   CDS   14.1207
FT   /*tag= a
FT   /product= Neural thread protein.
XX
PN   WO9615272-A1.
XX
PD   23-MAY-1996.
XX
PF   14-NOV-1995; 95WO-US17111.
XX
PR   14-NOV-1994; 94US-0340426.
XX
PA   (GEHO) GEN HOSPITAL CORP.
XX
PI   De LA MONTE S, Wands JR;

ALIGNMENTS

c	10	478.2	33.2	31853	22	AAI63343	Human kidney relat
c	11	477.4	33.1	11319	22	AAS39682	Genomic sequence #
c	12	477.4	33.1	11319	22	AAK90025	Human digestive sy
c	13	475.4	33.0	14796	19	AAV27941	Survivin gene. Ho
c	14	475.4	33.0	14796	22	AAS21523	DNA encoding human
c	15	475.4	33.0	14796	22	AAH47531	Human Her-3 genomi
c	16	471.2	32.7	12542	22	AAS26800	Human genomic DNA
c	17	467	32.4	15041	22	AAS35921	Human cardiovascul
18	466.8	32.4	9236	22	AAS03689	Rhesus gene locus:
c	19	466.8	32.4	9236	22	AAS03690	Rhesus gene locus:
c	20	466.8	32.4	9241	22	AAS03688	Rhesus gene locus:
21	461.8	32.0	12026	22	AAK67211	Human immune/haema
c	22	461.4	32.0	36221	22	AAS00624	Human death-associ
23	460.8	32.0	2368	22	AAK69566	Human digestive sy
24	459.4	31.9	30110	22	AAK89230	Human digestive sy
c	25	457.8	31.7	160552	22	AAD02697	Human glycosyl sul
26	456.4	31.7	38771	21	AAK81036	Human immune/haema
27	456.4	31.7	16310	21	AAF21086	Human low adenosin
28	456.4	31.7	16310	21	AAH34964	Human low adenosin
29	456.4	31.7	17634	21	AAF21087	Human adenosine re
30	456.4	31.7	17634	21	AAA34965	Human adenosine re
c	31	456.2	31.6	23130	22	AAK77045	Human immune/haena
32	455.2	31.6	44354	22	AAK77833	Human immune/haema
33	455.2	31.6	44354	22	AAK77836	Human immune/haema
34	455.2	31.6	44354	22	AAK77837	Human immune/haema
35	455.2	31.6	17245	22	AAK83897	Human immune/haena
36	455	31.6	161425	22	AAH02340	Human AKAP10 gene
37	455	31.6	162025	22	AAH02339	Human AKAP10 gene
38	455	31.6	13161	22	AAF97867	Human neuroblastom
39	452.6	31.4	4077	22	AAK69603	Human immune/haema
40	452.2	31.4	4077	22	AAK69605	Human immune/haema
41	452.2	31.4	10091	22	AAK69350	Human immune/haema
42	450	31.2	21693	22	AAK79930	Human immune/haema
c	43	449.4	31.2	1235	22	AAK87086	Human immune/haena
44	448.6	31.1	15745	22	AAK79250	Human immune/haema
45	448.4	31.1	16707	20	AAX58061	Genomic DNA for Hu

```
XX
DR    WPI; 1996-259865/26.
DR    P-PSDB; AAR95913.
XX
PT    Detection of neural thread protein in diagnosis of Alzheimer's
PT    disease - also NTP DNA and protein sequences used in gene and
PT    anti:sense therapy
XX
PS    Claim 24; Page 168-170; 238pp; English.
XX
CC    A method for detecting the presence of neural thread protein (NTP)
CC    having a molecular weight of 8, 14, 17, 21, 26 or 42 kD in a human
CC    subject comprises (a) contacting a sample from a human subject that
CC    is suspected of containing the NTP with at least one molecule
CC    capable of binding to the protein; and (b) detecting any of the
CC    molecule bound to the protein. The binding molecule is selected
CC    from an antibody free of natural impurities, a monoclonal antibody
CC    or a binding fragment of either of these. The method may be used for
CC    diagnosing the presence of Alzheimer's disease, neuroectodermal
CC    tumours and a malignant astrocytoma in a human.
XX
SQ    Sequence 1418 BP; 302 A; 396 C; 315 G; 405 T; 0 other;

      Query Match        84.8%;  Score 1223.4;  DB 17;  Length 1418;
      Best Local Similarity  96.9%;  Pred. No. 0;
      Matches 1375;  Conservative  0;  Mismatches  31;  Indels  13;  Gaps  12;
```

```
Qy      2 tttttttttgagatggagtttcacttctgttcgcccaggctggagtgcaatggcgcaat   61
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 tttttttttgagatggagtttcacttctgttcgcccaggctggagtgcaatggcgcaat   60

Qy     62 ctcagctcaccgcaacctccgcctcccgggttcaagcgattctcctgcctcagcctccc  121
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ctcagctcaccgcaacctccgcctcccgggttcaagcgattctcctgcctcagcctccc  120

Qy    122 agta-gctgggattacaggcatgtgccaccacgctcggctaattttgttattttttttag  180
          |||| |||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db    121 agtaggctgggattacaggcatgtgcca-ccacgctcggctaattttgttattttttttag  179

Qy    181 tagagatggagttctccatggttggtcaggctgtctcgaactcccgacctcagatgatc  240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    180 tagagatggagttctccatggttggtcaggctgtctcgaactcccgacctcagatgatc  238

Qy    241 cctccgctccggccctcccaaagtgctagatacagaggactgccacccatgcccg-ctctgc  299
          ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db    239 ctccgctccggccctcccaaagtgctagatacagaggactgccacccatgcccgctctgc  298

Qy    300 ctggctaattttgtggtagaaacagggtttcactgatgttgtgcccaggctggtctcctgag  359
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    299 ctggctaattttgtggtagaaacagggtttcactgatgttgtgcccaggctggtctcctgag  358

Qy    360 ctcaagcagtccacctgcctcagcctcccaaagtgctgggattacaggcgtgcagccgtg  419
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    359 ctcaagcagtccacctgcctcagcctcccaaagtgctgggattacaggcgtgcagccgtg  418

Qy    420 cctggcctttcttttattttttttaagacacaggtgtcccactcttaccccaggatgaa  479
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    419 cctggcctttcttttattttttttaagacacaggtgtcccactcttaccccaggatgaa  478

Qy    480 gtgcagtggtgtgatcacagctcactgcaagctcctgagatcaagacgcatcctct  539
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    479 gtgcagtggtgtgatcacagctcactgcaagctcctgagatcaagacgcatcctct  537

Qy    540 gcctcagcctccc-aagtagctgggaccaaagacatgcaccactacaacctggctaatttt  598
          ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db    538 gcctcagcctcccaaagtagctgggaccaaagacatgcaccactacaacctggctaatttt  597

Qy    599 tatttttattttttaatttttttgagacagagtctcactctgtcacccaggctggagtgca  658
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    598 tatttttattttttaatttttttgagacagagtctcactctgtcacccaggctggagtgca  657

Qy    659 gtggcgcaacttggctcactgcaacctccgcggttcaagttattctcctgcctcagcctcc  718
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    659 gtggcgcaacttggctcactgcaacctccgcggttcaagttattctcctgcctcagcctcc  718
```

```
Db    658 gtggcgcaacttggctcactgcaacctccgcctcccgggttcaagttattctcctgcctcc  717

Qy    719 cagcctcctgagtagctgggactacaggcgcgccaccacgctgcctcaattttttgtatt  778
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    718 cagcctcctgagtagctgggactacaggcgcgccaccacgctgcctcaattttttgtatt  777

Qy    779 tttagtagagatgggg-ttcaccatgttcgtccaggttgat-cttgatctctggaccttgt  836
          |||||||||||||||| ||||||||||||||||||||||||| ||||||||||||||||||
Db    778 tttagtagagatggggttcaccatgttccgccaggttgatgtgatctctggaccttgt  837

Qy    837 gatctgcctgcctcggcctccaaagtgctgggattacagg-cgtgagcacacgacccg  895
          |||||||||||||||||||||||||||||||||||||||| ||||| |||||||||||
Db    838 gatctgcctgcctcggcctccaaagtgctgggattacaggcgtgagcacacgacccg  897

Qy    896 gcttatttttaattttgtttgttttgaaatggaaatctcactctgttaccccaggctgagt  955
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    898 gcctattttaattttgtttgtttcgaaatggaaatctcactctgttaccccaggctgagt  957

Qy    956 gcaatggcgcaaatctcggctcactgcaacctctgcctcccgggctcaaggagattcctg  1015
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    958 gcaatggcgcaaatctcggctcactgcaacctctgcctcccgggctcaaggcgattcctg  1017

Qy   1016 tcctcagcctccaagcagctgggattacgggcacctgccaccacaccacccgctaatttg  1075
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1018 tcctcagcctccaagcagctgggattacgggcacctg-cgtgagcacacgccccaatttg  1076

Qy   1076 tattttcattagagacggggtttcaccatattgttcaggctggtctcaaactcctgacct  1135
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1077 tattttcattagagacggggtttcaccatattgttcaggctggtctcaaactctgacct  1135

Qy   1136 caggtgaccaactgcctcagcctcccaagtgctgggattgattgagcgccacctca  1195
          |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db   1137 caggtgaccaactgcctcagcctcccaagtgctgggattacaggcgtga--cgcctca  1194

Qy   1196 cccagcccggctaattttagtataaaaatatgtagacaatgggggtctttgctcttgccc  1255
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1195 cccagcccggctaattttagtataaaaatatgtagacaatgggggtctttgctcttgccc  1254

Qy   1256 aggctggtctcaaactcctggcttcatgcaatcctcccaaatgacacaacaccccagcc  1315
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1255 aggctggtctcaaactcctggcttcatgcaatcctcccaaatgacacaacaccccagcc  1314

Qy   1316 agtcacattttttaaacagttacatcttttattttagtactactagaaagtataacaataa  1375
          |||||||||||||||||||||||||||||||||||||||||||||| ||| |||||||||
Db   1315 agtcaca-ttttaaacagttacatcttttattttagtactactagaaagtgatacgataac  1373

Qy   1376 catgtcaaacctgcaaattcaagtagtaacagcagttcttt  1414
          ||||||||||||||||||||||||||||||||||||| |||
Db   1374 atgtcggaacctgcaaattcgagtagtacagagtctttt  1412
```

```
RESULT    2
AAQ77883
ID    AAQ77883 standard; cDNA; 1381 BP.
XX
AC    AAQ77883;
XX
DT    06-JUL-1995   (first entry)
XX
DE    Neural thread protein AD10-7 cDNA.
XX
KW    Neural thread protein AD10-7; Alzheimer's;
KW    malignant astrocytomas; glioblastomas; ss.
XX
OS    Rattus rattus.
XX
PN    WO9423756-A.
XX
PD    27-OCT-1994.
XX
PF    20-APR-1994;   94WO-US04321.
XX
```

```
PR   20-APR-1993;          93US-0050559.
XX
PA   (GEHO ) GEN HOSPITAL CORP.
XX
PI   De LA MONTE SM, Wands JR;
XX
DR   WPI; 1994-341497/42.
XX
PT   Detection of neural thread proteins - to detect sporadic and
PT   familial Alzheimer's disease, neuroectodermal tumours, malignant
PT   astrocytomas and glioblastomas (Eng)
XX
PS   Example 4; Fig 16R; 158pp; English.
XX
CC   AAQ77883 is the AD10-7 neural thread protein (NTP) cDNA. This
CC   sequence was used in the development of an antibody dependent
CC   method, for the detection of NTPs. This new method could be
CC   used to diagnose Alzheimer's disease (differentiating between
CC   sporadic and familial), neuroectodermal tumours, malignant
CC   astrocytomas and familial), neuroectodermal tumours, malignant
CC   astrocytomas and glioblastomas.
XX
SQ   Sequence 1381 BP; 296 A; 385 C; 301 G; 399 T; 0 other;

Query Match             74.8%;  Score 1078.6;  DB 15;  Length 1381;
Best Local Similarity   94.3%;  Pred. No. 1.1e-27;
Matches 1316;  Conservative    0;  Mismatches   54;  Indels   25;  Gaps   18;

Qy    2 tttttttttttgagagatggaggagttttcgctccttgttgtgcccaggctggagtgcagtggcgcaat 61
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 tttttttttttgagagatggaggagttttcgctccttgttgtgcccaggctggagtgcagtggcgcaat 60

Qy   62 ctcagctcaccgcaacctccgcctcccgggttcaagcgattctcctgcctcagcctccc 121
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 ctcagctcaccgcaacctccgcctcccgggttcaagcgattctcctgcctcagcctccc 120

Qy  122 agtagctcgggattacaggcatgtgcaccaccacgctcggctcaatttttgtattttttttagt 181
        |||||||||||||||||||||||||||||  |||||||||||| |||| ||||||||||||||||
Db  121 agtagctcgggattacaggcatgtgca--ccacgctcggctaatttgtattttttttttttagt 179

Qy  182 agagatggagtt--ctccaagtgttggtcaggctgtctcgaactccgacctcaggtgat 239
        ||||||||||||   ||||||| |||||||||||||||||||||||| |||||||| |||
Db  180 agagatggagttaactccaatgtggttcaggctgtctcgaactcccgacctcagatgat 239

Qy  240 ccctccgtctcggcctcccaaagtgct---agatacagacctgccaccaatgcccg-ct 295
        ||-| ||||||||||||||||||||||    || |||||||| |||||||||||||| ||
Db  240 ctcccgtctcggcctcccaaagtgctgagattacagacgtgagcaccaatgccggcct 299

Qy  296 ctgcctggctaattttgtggtagaaacaggggttcactgaty-tgcccaagtcggttctc 354
        ||||||||||||||||||||||||||||||||||||||||||  |||||||| ||||| |
Db  300 ctgcctggctaattttgtggtagaaacaggggttcactgatgttgcccaagctcggttcc 359

·Db  300 ctgcctggctaattttgtggtagaaacaggggttcactgatgttgcccaagctcggttcc 359

Qy  355 ctgagctcaagcaagtccaactgcctcagcctcccaaagtgctgggattacaggcgtgagccaccacg 414
        |||| |||||| |||||||||||||||||||||||||||||||||||||||||||  ||||||||| ||
Db  360 ctcagctcaagactgcctcagcctcccaaagtgctgggattacaggcgt-cag 418

Qy  415 ccgttgcctggccttctttatttttttaagacacaggtgtcccactcttaccagg 474
        |||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db  419 ccgttgcctggccttctttatttttttaagacacaggtgtaccactcttaccagg 478

Qy  475 atgaagtgcagtggtgtgatcacagctcgcgcctcaactcctgagatccagc-at 533
        |||||||||||||||||||||||||||||||||||||||||||||| |||||| ||
Db  479 atgaagtgcagtggtcgatcacagctcgcgcctcaactcctgagatccaagcaat 538

Qy  534 cctcctgcctcagcctcccaagtagctgggaccacacctacacctgcta 593
        |||||||||||||||||||||||||||||||||||||||||||||||||
Db  539 cctcctgcctcagcctcccaagtagctgggaccacacctacacctggta 597

Qy  594 atttttattttttagacagagtttcaccatgttgtcccaggctgga 653
        ||||||||||||| |||||||||||||||||||||||||||||||
Db  598 atttttatttttttaattttttgagacagagtctc-actgtcaccaggctgga 656

Qy  654 gtgcagtggcgcaatcttggctcactgcaacctccgggttcaagtgattattcctc 713
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  657 gtgcagtggcgcaatcttggctcactgcaacctctgcctcccgggttcaagttattctcc 716

Qy  714 tgcccagcctcctgagtagctgggactacaggggcccaccaccaccgcccggttaattttt 773
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  717 tgcccagcctcctgagtagctgggactacaggcgcccaccaccacgcccggcctaattttt 776

Qy  774 gtatttttagtagatggggtttcaccatgttggccaggttggtctcgatctcctggacc 832
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  777 gtatttttagtagagatgggtttcaccatgttgtccaggttgatcttctgacc 836

Qy  833 ttgtgatctgcctgcctcggcctcccaaagtgctgggattacaggcgtgagccaccacg 891
        ||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db  837 ttgtgatctgcctgcctcggcctcccaaagtgctgggattacag--gtcgtgactccac 894

Qy  892 cccgggcttattttgttttgttttgttttgtttttgaaatcctcactctgttacccaggctg 951
        |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db  895 gccggcctaattttttttttgttttgtttttgaaatcctcactctgttacccaggtcg 954

Qy  952 gagtgcaatggccaaatctcggctcactgcaacctctgcctcccgggctcaagcgattct 1011
        ||||||||||||| |||||||||||||||||||||||| |||||||||||||||||||
Db  955 gagtgcaatgg-caaatctcggctcactgcaacctctcgcctcccggg-tcaagcgattct 1012

Qy 1012 cctgtctcagcctcccaagcagctgggattacgggcgcacctgccaccaccacccgctaatt 1071
        |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db 1013 cctgtctcagcctcccaagcagctgggattacggg--acctgccacaccacgcctaatt 1070

Qy 1072 ttttgtatttttattagaggcgggtttcaccatatttggccaggct--ggtctcaaactcct 1130
        |||||||||||| |||||||||||||||||| |||| ||||||||  ||||||||||||||
Db 1071 tttgtatttttcattagagcgggtttt--accatattttcaggctggtctcaaactcct 1128

Qy 1131 gacctcaggtgatccagcccttccaaagtgctgggattacaggcgtgagcca 1190
        ||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db 1129 gacctcaggtgatccacccttccaaagtgctgggattacaggcgtgagcca 1188

Qy 1191 cctcaccagccgcggctaatttagataaaaataatatgtagcaatatggggtttcttgctgtg 1250
        ||||||||||||||||||||| ||||||||||||||||||||||||||||||||| |||||||
Db 1189 cctcaccagccggctaatttggaataaaaatatgtagcaatatggggtctttgctatgt 1246

Qy 1251 tgcccaggctggtctcaaactcctggcttcatgcaatccttccaaatgagccacaccacacc 1310
        |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db 1247 tgcccaggctggtctcaaactcctggcttcagtcaatccttccaaatgagccacaccacacc 1306

Qy 1311 cagccagtcacattttttaaacagttacatctttattttttttagtatactagaaagtaata 1370
        |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db 1307 cagccagtcacattttttaaacagttacatctttattttagtatactagaaagtaataca 1366

Qy 1371 ataaacatgtcaaac 1385
        |||||||||||||||
Db 1367 ataaacatgtcaaac 1381
```

RESULT   3
```
AAT27765
ID   AAT27765 standard; cDNA; 1381 BP.
XX
AC   AAT27765;
XX
DT   14-NOV-1996  (first entry)
XX
DE   AD 10-7 human neural thread protein clone (partial sequence).
XX
KW   Neural thread protein; NTP; diagnosis; detection;
KW   Alzheimer's disease; neuroectodermal tumour; malignant astrocytoma;
KW   monoclonal antibody; binding fragment; ds.
XX
OS   Homo sapiens.
XX
PN   WO9615272-A1.
XX
PD   23-MAY-1996.
XX
PF   14-NOV-1995;   95WO-US17111.
```

XX
PR   14-NOV-1994;   `94US-0340426.
XX
PA   (GEHO) GEN HOSPITAL CORP.
XX
PI   De LA MONTE S, Wands JR;
XX
DR   WPI; 1996-259865/26.
XX
PT   Detection of neural thread protein in diagnosis of Alzheimer's
PT   disease - also NTP DNA and protein sequences used in gene and
PT   anti:sense therapy
XX
PS   Example 4c; Figure 16R; 238pp; English.
XX
CC   A method for detecting the presence of neural thread protein (NTP)
CC   having a molecular weight of 8, 14, 17, 21, 26 or 42 kD in a human
CC   subject comprises (a) contacting a sample from a human subject that
CC   is suspected of containing the NTP with at least one molecule
CC   capable of binding to the protein; and (b) detecting any of the
CC   molecule bound to the protein. The binding molecule is selected
CC   from an antibody free of natural impurities, a monoclonal antibody
CC   or a binding fragment of either of these. The method may be used for
CC   diagnosing the presence of Alzheimer's disease, neuroectodermal
CC   tumours and a malignant astrocytoma in a human. A number of clones
CC   of neural thread protein were isolated from healthy 17-18 week old
CC   foetal human brain (HB) 2 year old temporal lobe neocortex and end
CC   stage Alzheimer's disease (AD) cerebral cortex. See AAT27753-75.
XX
SQ   Sequence 1381 BP; 385 C; 301 G; 400 T; 0 other;

Query Match   74.7%; Score 1077; DB 17; Length 1381;
Best Local Similarity   94.3%; Pred. No. 3e-277;
Matches 1315; Conservative   0; Mismatches 55; Indels 25; Gaps 18;

Qy   2 ttttttttttgagatggagtttcgctcttgttgcccaggctggagtgcaatggcgcaat 61
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ttttttttttgagatggagtttcgctcttgttgcccaggctggagtgcaatggcgcaat 60

Qy   62 ctcagctcaccgcaacctccgcctcccggttcaagcgattctcctgcctcagcctccc 121
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 ctcagctcaccgcaacctccgcctcccggttcaagcgattctcctgcctcagcctccc 120

Qy   122 agtagctcggattacaggcatgtgccaccacgcccggcccagctaatttttgtatttttagt 181
            |||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db   121 agtagctcggattacaggcatgtgca--ccacgcccggcctaatttttgtattttttagt 179

Qy   182 agagatggggttt--ctccatgtgtcaggctggtctctcgaactcccgacctcaagtgat 239
            |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db   180 agagatggggtttaacctccatgtgtcaggctggtctctcgaactcccgacctcagatgat 239

Qy   240 ccctccgtctccggcctcccaaagtgct--agatacaggactggcgccaccatgcccgc--ct 295
            |||||||||||||||||||||||| ||||||||||||||||||||||||||| ||
Db   240 ctcccgtctccggcctcccaaagtgctgagatacaggcatgagcgaccaccatgcccggcct 299

Qy   296 ctgcgcctggctaatttttgtggtagaaacaggggttcactggtt-tgcccaagtggttctc 354
            ||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db   300 ctgcgcctggctaatttttgtggtagaaacaggggttcactgattgtgcccaagctggtctc 359

Qy   355 ctgagctcaagcaagtcacctgcctcagcctcccaaagtgctggggattacaggcgtgcag 414
            ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   360 ctgacctcaagcaagtcacctgcctcccaaagtgctggggattacaggcgt-cag 418

Qy   415 ccgtgcctggccctttattttttttaagacacaggtgtccacctcttacccagg 474
            || |||||||||||||||||||||||||||||||||||||||||||||||||
Db   419 ccgtgcctggccctttattttttttaagacacaggtgtaccacctcttacccagg 478

Qy   475 atgaagtcgcagtggttgtgatcacagcctcactcctgagatcaaagc--at 533
            ||||||||||||||||||||||||||||||||||||||||||||||| ||
Db   479 atgaagtcgcagtggttgtgatcacagctcactgcacctcctgagatcaaagcaat 538

Qy   534 cctcctgcctcagctcccaagtagctcgggaccaaagacaccatgcaccactacccactggcta 593
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   539 cctcctgcctcagctcccaagtagctgggaccaaagacaccatgcaccactacccactgg-ta 597

Qy   594 attttaatttttattttattttgagacagagtctcaaactgctgtcacccaggctgga 653
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   598 attttaatttttattttattttgagacagagtctc-actctgtcacccaggctgga 656

Qy   654 gtgcaagtggtgcgcaatcttggctcactgcaacctctgcctcccgggttcaagttattctcc 713
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   657 gtgcagtggcgcaatcttggctcactgcaacctctgcctcccgggttcaagttattctcc 716

Qy   714 tgcccagcctcctgagtagctgggactacaggcgcccaccacgcccagctaatttttt 773
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   717 tgcccagcctcctgagtagctgggactacaggcgcccaccacgcctagctaatttttt 776

Qy   774 gtatttttagtagagatgggg--ttcacctgttccgccaggttgatcttgatcctctggacc 832
            ||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db   777 gtatttttagtagagatggggttcacctgttccgccaggttgatcttgatcctctggacc 836

Qy   833 tttgtgatctgcctgcctcggcct-cccaaagtgctgggattacaggcgtgagccaccacg 891
            ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db   837 tttgtgatctgcctcggcctcccaaagtgctgggattacag--gtcgtgactcac 894

Qy   892 ccccggcttatttttgttttgttttgaattgaatggaatctcactctgttaccaggctg 951
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   895 gccggccttatttttgttttgttttgaattgaatggaatctcactctgttaccaggtcg 954

Qy   952 gagtgcaatggccaaatctcggctcactgcaacctctgcctcccgggttcaaggggatcct 1011
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   955 gagtgcaatgg-caaatctcggctactgcaacctctgcctcccgggt-tcaagggatcctc 1012

Qy   1012 cctgtctcagccttcctccaagccaagcaggttggattacggtgcacctgcacacccgctaatt 1071
            ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db   1013 cctgtctcagcctcccaagccaagcagctgggattacggg--acctgcacacacccgctaatt 1070

Qy   1072 tttgttatttcattagaggcggggtttcaccatgttgtccaggct--ggtctcaaactcct 1130
            |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db   1071 tttgttatttcattagaggcggggtttcaggct---accatattgtgtccaggttgtctcaaactcct 1128

Qy   1131 gacctcaggtgacccacctgcctccagccttccaaagtgctgggattacaggcgtgagcca 1190
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1129 gacctcaggtgacccacctgcctcagccttccaaagtgctgggattacaggcgtgagcca 1188

Qy   1191 cctcacccagccgcgcctaatttttttgtatttttagtagtacagatacatgtgtgttctgtgt 1250
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1189 cctcacccagccgcctaatttttgttataaaaatatgtagcaatggggggttctgtgt 1246

Qy   1251 tgcccaggctggtctcaaacttctgggctcatgcaatcctcccaaatgagccacacacc 1310
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1247 tgcccaggctggtctcaaacttctggcctcaagtcatcctcccaaatagagccacacacc 1306

Qy   1311 cagccagctcacatttttaaacacagttacatctttatttgtagtataactataaagtaataca 1370
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1307 cagccagtcacattttttaaacagttaacatcttatttatttagtatactaaaagtaataca 1366

Qy   1371 ataaacatgtcaaac 1385
            |||||||||||||||
Db   1367 ataaacatgtcaaac 1381

RESULT 4
AAK69279/c
ID   AAK69279 standard; DNA; 33747 BP.
XX
AC   AAK69279;
XX
DT   06-NOV-2001   (first entry)
XX
DE   Human immune/haematopoietic antigen genomic sequence SEQ ID NO:24091.
XX
KW   Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW   cytostatic; gene therapy; vaccine; metastasis; ds.
XX

```
OS   Homo sapiens.
XX
PN   WO20015718Z-A2.
XX
PD   09-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01354.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   14-JUL-2000; 2000US-0218290.
PR   11-JUL-2000; 2000US-0217496.
PR   11-JUL-2000; 2000US-0217487.
PR   07-JUL-2000; 2000US-0216880.
PR   07-JUL-2000; 2000US-0216647.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   23-AUG-2000; 2000US-0227182.
PR   30-AUG-2000; 2000US-0227009.
PR   01-SEP-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   05-SEP-2000; 2000US-0229945.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   08-SEP-2000; 2000US-0231414.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0232080.
PR   08-SEP-2000; 2000US-0232081.
PR   12-SEP-2000; 2000US-0231968.
PR   14-SEP-2000; 2000US-0232397.
PR   14-SEP-2000; 2000US-0232398.
PR   14-SEP-2000; 2000US-0232399.
PR   14-SEP-2000; 2000US-0232400.
PR   14-SEP-2000; 2000US-0232401.
PR   14-SEP-2000; 2000US-0233063.
PR   14-SEP-2000; 2000US-0233064.
PR   14-SEP-2000; 2000US-0233065.
PR   21-SEP-2000; 2000US-0234223.
PR   14-SEP-2000; 2000US-0234274.
PR   25-SEP-2000; 2000US-0234997.
PR   25-SEP-2000; 2000US-0234998.
PR   26-SEP-2000; 2000US-0235484.

PR   27-SEP-2000; 2000US-0235834.
PR   27-SEP-2000; 2000US-0235836.
PR   29-SEP-2000; 2000US-0236327.
PR   29-SEP-2000; 2000US-0236367.
PR   29-SEP-2000; 2000US-0236368.
PR   29-SEP-2000; 2000US-0236369.
PR   29-SEP-2000; 2000US-0236370.
PR   02-OCT-2000; 2000US-0237037.
PR   02-OCT-2000; 2000US-0237038.
PR   02-OCT-2000; 2000US-0237039.
PR   02-OCT-2000; 2000US-0237040.
PR   13-OCT-2000; 2000US-0239935.
PR   13-OCT-2000; 2000US-0239937.
PR   20-SEP-2000; 2000US-0240960.
PR   20-OCT-2000; 2000US-0241221.
PR   20-OCT-2000; 2000US-0241786.
PR   20-OCT-2000; 2000US-0241785.
PR   20-OCT-2000; 2000US-0241787.
PR   20-OCT-2000; 2000US-0241808.
PR   20-OCT-2000; 2000US-0241809.
PR   20-OCT-2000; 2000US-0241826.
PR   01-NOV-2000; 2000US-0244617.
PR   08-NOV-2000; 2000US-0244474.
PR   08-NOV-2000; 2000US-0246475.
PR   08-NOV-2000; 2000US-0246476.
PR   08-NOV-2000; 2000US-0246477.
PR   08-NOV-2000; 2000US-0246478.
PR   08-NOV-2000; 2000US-0246523.
PR   08-NOV-2000; 2000US-0246524.
PR   08-NOV-2000; 2000US-0246525.
PR   08-NOV-2000; 2000US-0246526.
PR   08-NOV-2000; 2000US-0246527.
PR   08-NOV-2000; 2000US-0246528.
PR   08-NOV-2000; 2000US-0246532.
PR   08-NOV-2000; 2000US-0246609.
PR   08-NOV-2000; 2000US-0246610.
PR   08-NOV-2000; 2000US-0246611.
PR   17-NOV-2000; 2000US-0246613.
PR   17-NOV-2000; 2000US-0249207.
PR   17-NOV-2000; 2000US-0249208.
PR   17-NOV-2000; 2000US-0249209.
PR   17-NOV-2000; 2000US-0249210.
PR   17-NOV-2000; 2000US-0249211.
PR   17-NOV-2000; 2000US-0249212.
PR   17-NOV-2000; 2000US-0249213.
PR   17-NOV-2000; 2000US-0249214.
PR   17-NOV-2000; 2000US-0249215.
PR   17-NOV-2000; 2000US-0249216.
PR   17-NOV-2000; 2000US-0249217.
PR   17-NOV-2000; 2000US-0249218.
PR   17-NOV-2000; 2000US-0249244.
PR   17-NOV-2000; 2000US-0249245.
PR   17-NOV-2000; 2000US-0249264.
PR   17-NOV-2000; 2000US-0249265.
PR   17-NOV-2000; 2000US-0249297.
PR   17-NOV-2000; 2000US-0249299.
PR   17-NOV-2000; 2000US-0249300.
PR   01-DEC-2000; 2000US-0250160.
PR   01-DEC-2000; 2000US-0250391.
PR   05-DEC-2000; 2000US-0251030.
PR   05-DEC-2000; 2000US-0251988.
PR   06-DEC-2000; 2000US-0256719.
PR   05-DEC-2000; 2000US-0251479.
PR   07-DEC-2000; 2000US-0251856.
PR   08-DEC-2000; 2000US-0251868.
PR   08-DEC-2000; 2000US-0251869.
PR   08-DEC-2000; 2000US-0251989.
PR   08-DEC-2000; 2000US-0251990.
PR   11-DEC-2000; 2000US-0254097.
PR   05-JAN-2001; 2001US-0259678.

XX
PA   (HUMA-) HUMAN GENOME SCI INC.
```

```
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-483426/52.
XX
PT   Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT   useful for preventing, diagnosing and/or treating cancers and
PT   metastasis -
XX
PS   Disclosure; SEQ ID NO 24091; 307lpp + Sequence Listing; English.
XX
CC   AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC   amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC   activity, and can be used in gene therapy and vaccine production.(I)
CC   proteins and polynucleotides may be used in the prevention, diagnosis and
CC   treatment of diseases associated with inappropriate (I) expression. For
CC   example, they may be used to treat disorders associated with decreased
CC   expression by rectifying mutations or deletions in a patient's genome
CC   that affect the activity of (I) by expressing inactive proteins or to
-CC  supplement the patients own production of (I). Additionally, (I)
CC   polynucleotides may be used to produce the secreted (I), by inserting
CC   the nucleic acids into a host cell and culturing the cell to express the
CC   protein. (I) proteins and polynucleotides may be used to prevent,
-CC  diagnose and treat immune/haematopoietic-related diseases, especially
CC   cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC   to AAK87694 represent human immune/haematopoietic antigen genomic
CC   sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC   represent sequences used in the exemplification of the present invention.
XX
SQ   Sequence 33747 BP; 9625 A; 7893 C; 7691 G; 8538 T; 0 other;
```

Query Match   35.9%; Score 517.2; DB 22; Length 33747;
Best Local Similarity   68.8%; Pred. No. 2.1e-127;
Matches 844; Conservative   0; Mismatches 363; Indels 19; Gaps   9;

```
Qy      1 ttttttttttttgagatggaagtttccgctcttgtctgcccaggctggagtgcaatggcgcaa 60
          ||||||||||||||||||||||||||||||||||||||||  | || ||||||| |||||||||
Db  33207 TTTTTTTTTTTTTTGAGATGGGAGTCT--CGCTGCTGCCCAGGTTGGAGTGCAGTGGTGA 33150

Qy     61 tctcagctcaccgcaccactccgcctccccgggttcaagcgattctcctgcctcagcctcc 120
          ||||||||||||||||| ||||||||  |||||||||||||| ||||||||||||||||||
Db  33149 TCTCAGCTCACTGCAACCTCTGCCTCCCCAGGTTTTAAGCGATTCCTCGACTTAGCCTCCC 33090

Qy    121 cagtagctgtgggattacaggcatgtgcaccacgctcggctaatttttgtatttttttag 180
          ||||||||||||| |||||||||||||||||||||||| ||||||||| || |||||||||
Db  33089 TAGAAGCTGTGGGATTATAGGCCTCGCCACCACATGCCTGGCTAATTTT---TGTATTTTTAG 33033

Qy    181 tagagatggagtttctccatgttggtcaggctggtctcgaactcccgacctcaagtgatc 240
          ||||||||||||||||||||||| |||||||||||||||||||||||| ||||||||||
Db  33032 CAGAGACGGGGTTTTGCCATGTTGGCCAGGCTGGTCTCAAACTCCTGAACTCAGGTGATC 32973

Qy    241 cctccgtctcggcctcccaaaagtgctagatacagggctggcacccgggctctgcc 300
          ||||| ||||||||||||||||||||||||||||||||| ||||||| | || || |
Db  32972 TGCCTGCCTTGGCCTCCCAAAGTGCTGGGATTACAGTTGTGAGCCACTGCACCGGCCTA 32913

Qy    301 tggctaattttttgtggtagaaacagggtttcactgatgtgcccaagctggtctcc----- 355
          ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  32912 TTTTTTTTATTATTTTTTTGTTGAGACAGGAGTCTCACTCTGTCTTTTGTCCAGGCTGGAGTGCAGTGG 32853

Qy    356 tgagctcaagcaagtccaactcgcctccaagtgctgggattacaggcgtgcagc 415
          ||| |||||||||||||||||||||| || |||||||||||| |||||||||||
Db  32852 TATGATCACAGCTCATTGCACCTCCAACTCCTGGACTCAGGCAATCCTCCACCTTAGC 32793

Qy    416 cgtgcctggcctttttattttttaagacacaggtgtcccactcttacccagga 475
          | |||||||||||| |||||||| ||||||||||| |||||||||||||||||
Db  32792 TGCCTGAGTAGCTGGGACTACAACAAGCTTGGCGCCACCATGTCTGGCTAATTTTTTGTATTTTT 32733

Qy    476 tgaagtgcagtggtgtgatca--cagctcactgcagcctccaactcctgagatcaagc-a 532
          ||||||||||||||| ||||  ||| |||| |||||||||||||||||||||||| ||
Db  32732 TGTAGAGATGGGGTTTTGCAGCTGTTCAGCTAGTGTTGAACTCCTGGCCTCAAGCAA 32673

Qy    533 tcctcctgcctcagcctcccaagtagctgggaccaaagacatgcaccactacaacctggct 592
          ||||| |||||||||||||| ||| || |||| ||||| ||||||||||||| ||||||||
```

```
Db  32672 TCCTCCCCGCCCTTGGTTGGGTTACCAAAGTGTTGGGGATTATAGGCGATTAGCCGATTCTC 32613

Qy    593 aatttttattttttattttttaattttttgagacagagtcctcaactctgtcaccccaggctgg 652
          || |||| |||||| |||||| |||||| |||||||||| ||| |||||| | ||||| ||||
Db  32612 TATACAT-TGGCTTTTTTTTTTTTTTTTTTAGATGGAGTCTC-ACTCTGTTGGCCAGGCTGG 32555

Qy    653 agtgcagttggcgcaatcttggctcactgcaacctctgcctcccgggttcaagttattctc 712
          ||| |||| ||| |||| |||||| |||||| ||| |||||||| ||||||||  |||||||
Db  32554 AGTGCAGTGGGGCGCAATCTTGGCTCACTGCTTACTGTGAACCTCTGCCTCCCGGGTTCAAGCGATTCTC 32495

Qy    713 ctgcccaacctcctgagtagctgggactacaggcgc---ccacaacgcctagctaattt 769
          ||||||||||||||||| || |||||||||||| |    |||| ||| |||||||||||
Db  32494 CTGCCTCAGCCTCCTTAAGTAGCTGGGACTATAGGCGCCCACCACACCCGGGCCCAATTT 32435

Qy    770 ttttgtattttttagtagagacgggttccaccatgttccaggttgatcttgatctctgg 829
          |||||||||||||||||||||||||||  |||| |||||||  ||||||||||||| |||
Db  32434 TTTGTATTTTTTAGTAGAGACATGGGGTTTCACCATGTTGGCCAGGCTGGTCTCGAACTTCTG 32375

Qy    830 aacttgtgatcctgcctcggcctcccaaaagtgctgggattacaggcgtgggtctcgaacttctg 889
          ||||||| || ||||| |||||||||||||||||||||||||| ||  ||| ||| ||||
Db  32374 ACCTTGTGATCTGTCTGCCCTGGCCTCCCAAAATGCTGGGATTACAGGCGATGAGCCACTG 32315

Qy    890 cgcccggcttattttaattttgttttgaaatggaatccactctgttaccoagc 949
          ||||||| |||||||||||||| || |||||||| ||| |||||||||| |||
Db  32314 CACCCGGGCCGGCACTTTT-TTTTTCTTGGAGACGGGAGTCTCACTCACTTTGTCGCCCAGGC 32256

Qy    950 tggagtgcaatggcgcaaatctcggctcactgcaacctctgcctcccgggctcaagcgatt 1009
          ||| ||| ||||| |||||||| |||||||||||||||||||||||||||| ||| |||||
Db  32255 TGGAGTACAGTGGCATGGATCTCAGCTCACTGCAACCTCCACTCCCCTGGGTTCGAGTCATT 32196

Qy   1010 ctcctgtctcagcctccaagcagctgggattacaggcgccacacacccgctaa 1069
          |||||||||||||||||||||||| |||||||||||| |||||| ||||||||
Db  32195 TCCCTGCCTCANCCTCCCCCAGTNAGCTGGGACTACAGGCGTATGCCACCATACCCGTCTAA 32136

Qy   1070 tttttgtattttcattagagggggttcaccatattttgtcaggctggtctcaaactcc 1129
          |||||| |||| ||||| ||||||||||||||||| ||| |||||||||||||||||||
Db  32135 TTTTTGTNGTTTTTTAGTTGAGATGGGAGTTTTCACCATGTTGGCCAGGCTGGTCTCCAACTCC 32076

Qy   1130 tgacctgaggtgaccoaacctgcctcagcctcccaaagtgctgggattacaggcgtgagcc 1189
          |||||||||| ||| ||||||||||||||||||||||||||||||||||| |||||| |||
Db  32075 TGACCTTAAGTGATCTGCCCGCCCGCCTCCAAAGTGCTGGGATTACAGGCCATGAGCC 32016

Qy   1190 acctcaccagccgcggctaatttagat 1215
          |||||||||||| |||||||||||| |
Db  32015 ACTGTGCTCGGCCCCATTGACTTTTTAT 31990
```

```
RESULT   5
AAK73093
ID   AAK73093 standard; DNA; 33747 BP.
XX
AC   AAK73093;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human immune/haematopoietic antigen genomic sequence SEQ ID NO:27905.
KW   Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW   cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS   Homo sapiens.
XX
PN   WO200157182-A2.
XX
PD   09-AUG-2001.
XX
PF   17-JAN-2000; 2001WO-US01354.
XX
PR   31-JAN-2000; 200000-0179065.
PR   04-FEB-2000; 200000-0180628.
PR   24-FEB-2000; 200000-0184664.
PR   02-MAR-2000; 200000-0186350.
```

PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0211135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216080.
PR 11-JUL-2000; 2000US-0217487.
PR 14-JUL-2000; 2000US-0217496.
PR 26-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 14-AUG-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 18-AUG-2000; 2000US-0225759.
PR 14-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
-PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 08-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0232338.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 16-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 14-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 26-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 27-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.

PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241825.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244616.
PR 08-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 01-NOV-2000; 2000US-0249244.
PR 01-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 01-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251856.
PR 05-DEC-2000; 2000US-0251868.
PR 06-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 08-DEC-2000; 2000US-0254097.
PR 11-DEC-2000; 2001US-0256719.
PR 05-JAN-2001; 2001US-0259678.

XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA,   Barash SC,   Ruben SM;
XX
DR WPI; 2001-483426/52.
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
    metastasis -
XX
PS Disclosure; SEQ ID NO 27905; 3071pp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I

```
CC   amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC   activity, and can be used in gene therapy and vaccine production. (I)
CC   proteins and polynucleotides may be used in the prevention, diagnosis and
CC   treatment of diseases associated with inappropriate (I) expression. For
CC   example, they may be used to treat disorders associated with decreased
CC   expression by rectifying mutations or deletions in a patient's genome
CC   that affect the activity of (I) by expressing inactive proteins or to
CC   supplement the patients own production of (I). Additionally, (I)
CC   polynucleotides may be used to produce the secreted (I), by inserting
CC   the nucleic acids into a host cell and culturing the cell to express the
CC   protein. (I) proteins and polynucleotides may be used to prevent,
CC   diagnose and treat immune/haematopoietic-related diseases, especially
CC   cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC   to AAK87694 represent human immune/haematopoietic antigen genomic
CC   sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC   represent sequences used in the exemplification of the present invention.
XX
SQ   Sequence 33747 BP; 8538 A; 7691 C; 7893 G; 9625 T; 0 other;
```

```
Query Match    35.9%;  Score 517.2;  DB 22;  Length 33747;
Best Local Similarity   68.8%;
Matches  844;  Conservative    0;  Mismatches  363;  Indels   19;  Gaps    9;

Qy      1 tttttttttttttgagatggagttcgctcttgttgcccaggctggagtgcaatggcgcaa 60
          ||||||||||||| ||||||||||||    |||| |||||||||||||||        ||
Db    541 tttttttttttttgagatggagtct--cgctctgtgcccaaggctggagtcagtgatgtga 598

Qy     61 tctcagctcaccgcaacctccgcctcccggggttcaagcgagttctcctgcctcagcctcc 120
          ||||||||||| |||||||| ||||||| ||| ||||||| |||||  | |||||||||||
Db    599 tctcagctcactgcaacctctgcctcccaggttaagcgagatcctcccggcctcagcctcc 658

Qy    121 cagtagctgggattacaggcgcatgtgcaccacctcggcacgctcggcctaattttgtatt 180
          | ||||||||||||| |||||||     ||||| |||||   ||||| ||||||  |||||
Db    659 tagaagctgggattataggcgcacaccaccaccgctcggctaattttttgtatttttag 715

Qy    181 tagaggatggagtttctccaatgttgggctggtctcgaactcccgacctcaggatgatc 240
          |||| | ||| ||||| |||||||||| ||||||||| ||||| ||||||||| |||||
Db    716 cagagacgggtttttgccatgttggccaggctggtctcaaactcctgaacctcaggtgatc 775

Qy    241 cctccgtctcggcctcccaaaagtgctagatacaagcagccacactgcccggctctgcc 300
          | || |||| ||||||||||| ||||| ||| | | |||||||||||| |||||||| ||
Db    776 tgcctgccttggcctcccaaagtgctgggattacagttgtgagccactgcgcccggcta 835

Qy    301 tggctaattttgttggtagaaacaggtgttcactgatgtgccaagctggttctcc---- 355
          ||||||||| ||||| |||| |||| |||| ||| | ||| |||||||| |||
Db    836 tttttattatttgttgagacagagtctcttttgtccaggctggagtgcagtg 895

Qy    356 tgagctcaagcaagtccacctgcctcagcctcccaagtgctgggattacaggcgtgagc 415
          || | ||||| || | | || |||||| |||||||||||||||||||||||  ||||||
Db    896 tatgatacaagctcattgcagcctccaaactcctgactacagcaatcctcccaaccttagc 955

Qy    416 cgtgcctggccttttatttttaagacacaggtgtcccactcttaccaggga 475
          |  | ||| || || ||||||| | || | |  ||||||||||||||| |
Db    956 tgcctgagtcagctagctggggactacaagcttcgcccacactgctggctaatttttgtatttt 1015

Qy    476 tgaagtgcagtggtgtgatca--cagctcactgcagcctccaaactcctgagatcaagc-a 532
          ||  || ||| |||||| | | |||||||| ||| |||||||||||||||| |||| ||
Db   1016 tgtagagatggggtttcgccatgctagtcgttgaactcctgggctcaagcaa 1075

Qy    533 tcctcctgcctcagcctcccaagtagctgggaccaaagacatgcaccaccacctgct 592
          |||||||||||||||||||||||| |||||||  | | | | || ||||| ||||
Db   1076 tcctcccgccttggcctcccaaagtgtgggattacaggcgtgagccattgccggacc 1135

Qy    593 aatttttattttttaatttttttttggagacagagtctcactctgtcacccaggctg 652
          |||||||||||||| | ||||||||||||||| || ||||| |||| ||||||||||
Db   1136 tatacat-tggcttttttttttagatggagtctc-actcgttggccaggctg 1193

Qy    653 agtgcagtggcgcaatcctggctcactgcaacctctgcctcccgggttcaagttattctc 712
          |||| ||||| ||||  ||||||||||||||||| ||||||| |||||||||  |||||
Db   1194 agtgcagtggcgcaatctggctcactgcaacctctcctgggttcaagcgattctc 1253

Qy    713 ctgcccccagcctcctgagtagctgggactacaggcgc---ccaccacgccagctaattt 769
          |||||||| ||||| |||||||||||||||||||||||     ||||||| ||||||||||
Db   1733 actgtgctcggccattgacttttat 1758
```

```
Db   1254 ctgcctccagcctcctaagtagctcggggactataggcgcatcactccggtccaattt 1313
             |||||||||||      |||||||||||||    |||||||||  || ||  ||||||
Qy    770 tttgtattttaagtagagatggggtttcaccatgttcgtccaggctgatctctctctgg 829

Db   1314 tgtattttaagtagagatggggtttcaccatggtccaggctgtctcgaactcg 1373
           | ||||||||||   ||||||||||  |||||    ||||||  |  | ||| |||
Qy    830 acttgtgatctgcctgcctcggcctcccaaagtgctgggattacaggcgtgagccacca 889

Db   1374 aacttgtgatctgcctggcctcccaaaatgctgggattacaggcgtgagccactg 1433
           |||||||||||||||| || ||||||| ||||||||||||||||||||| |||| |
Qy    890 cgcccggcttatttttaattttgttttgttgtggaatctcactcatgttaccaggc 949

Db   1434 cacccggccgactttt--ttttttcctggagacggagtctcactcaacttgtcgcccaggc 1492
           |||||||| |||||||   ||||| |||| |||||||||||||| | |||| | |||||||
Qy    950 tggagtgcaaatctcggctcactgcaacctctgcctcccgggttcaagcgattacaggcat 1009

Db   1493 tggagtcagtggattataggcgcaccacaccacgctcggctaacttcctgggttcggtcatt 1552
           |||||| |||| ||||||||||||||| ||||||||||||||  | |||||||| | ||||
Qy   1010 ctcctgtctcagcctcccaagcagctgggattacaggccaccaccacccgctaa 1069

Db   1553 tcctgcctcatccccagtagctgcgattacaggtatgcaccaccatacccggctaa 1612
           ||| |||| || || |||||||||  ||||||  |||| ||| || | ||| ||||
Qy   1070 tttttttgtattttagagcgggtttcaccatcattgtcaggctgctcaaactcc 1129

Db   1613 ttttttgtgttttagttgagatggggtttcaccatgttagctgggtaagctggtcctcaaactcc 1672
           |||||||| ||||||| |||||||||||||||||||||  |||| ||| ||||| ||||||||||
Qy   1130 tgacctcaggtgacccacctgcctcagcctcccaaagtgctgggattacaggcgtgagcc 1189

Db   1673 tgacctcaagtgatctgcccgcctcagcctcccaaagtgctggggattacaggcatgagcc 1732
           |||||||| ||||| ||||  |||||||||||||||||||||| ||||||||||| |||||
Qy   1190 acctcaccagccggctaatttagat 1215

Db   1733 actgtgctcggccattgacttttat 1758
```

```
PR   26-JUL-2000;  2000US-0220963.
PR   26-JUL-2000;  2000US-0220964.
PR   14-AUG-2000;  2000US-0244518.
PR   14-AUG-2000;  2000US-0244519.
PR   14-AUG-2000;  2000US-0245213.
PR   14-AUG-2000;  2000US-0245214.
PR   14-AUG-2000;  2000US-0245266.
PR   14-AUG-2000;  2000US-0245267.
PR   14-AUG-2000;  2000US-0245268.
PR   14-AUG-2000;  2000US-0255270.
PR   14-AUG-2000;  2000US-0255447.
PR   14-AUG-2000;  2000US-0255757.
PR   14-AUG-2000;  2000US-0255758.
PR   14-AUG-2000;  2000US-0255759.
PR   18-AUG-2000;  2000US-0226279.
PR   22-AUG-2000;  2000US-0226681.
PR   22-AUG-2000;  2000US-0226868.
PR   22-AUG-2000;  2000US-0227182.
PR   23-AUG-2000;  2000US-0227009.
PR   30-AUG-2000;  2000US-0228924.
PR   01-SEP-2000;  2000US-0229287.
PR   01-SEP-2000;  2000US-0292287.
PR   01-SEP-2000;  2000US-0229343.
PR   01-SEP-2000;  2000US-0229344.
PR   01-SEP-2000;  2000US-0229345.
PR   05-SEP-2000;  2000US-0229509.
PR   05-SEP-2000;  2000US-0229513.
PR   06-SEP-2000;  2000US-0230437.
PR   06-SEP-2000;  2000US-0230438.
PR   08-SEP-2000;  2000US-0231242.
PR   08-SEP-2000;  2000US-0231243.
PR   08-SEP-2000;  2000US-0231244.
PR   08-SEP-2000;  2000US-0231413.
PR   08-SEP-2000;  2000US-0231414.
PR   08-SEP-2000;  2000US-0232080.
PR   08-SEP-2000;  2000US-0232081.
PR   12-SEP-2000;  2000US-0231968.
PR   14-SEP-2000;  2000US-0232397.
PR   14-SEP-2000;  2000US-0232398.
PR   14-SEP-2000;  2000US-0232399.
PR   14-SEP-2000;  2000US-0232400.
PR   14-SEP-2000;  2000US-0232401.
PR   14-SEP-2000;  2000US-0233063.
PR   14-SEP-2000;  2000US-0233064.
PR   14-SEP-2000;  2000US-0233065.
PR   21-SEP-2000;  2000US-0234223.
PR   21-SEP-2000;  2000US-0234274.
PR   25-SEP-2000;  2000US-0234997.
PR   25-SEP-2000;  2000US-0234998.
PR   26-SEP-2000;  2000US-0235484.
PR   27-SEP-2000;  2000US-0235836.
PR   29-SEP-2000;  2000US-0236327.
PR   29-SEP-2000;  2000US-0236367.
PR   29-SEP-2000;  2000US-0236368.
PR   29-SEP-2000;  2000US-0236369.
PR   29-SEP-2000;  2000US-0236370.
PR   02-OCT-2000;  2000US-0236802.
PR   02-OCT-2000;  2000US-0237037.
PR   02-OCT-2000;  2000US-0237038.
PR   02-OCT-2000;  2000US-0237039.
PR   02-OCT-2000;  2000US-0237040.
PR   13-OCT-2000;  2000US-0239935.
PR   13-OCT-2000;  2000US-0239937.
PR   20-OCT-2000;  2000US-0240960.
PR   20-OCT-2000;  2000US-0241221.
PR   20-OCT-2000;  2000US-0241785.
PR   20-OCT-2000;  2000US-0241786.
PR   20-OCT-2000;  2000US-0241787.
PR   20-OCT-2000;  2000US-0241808.
PR   20-OCT-2000;  2000US-0241809.
PR   20-OCT-2000;  2000US-0241826.
PR   01-NOV-2000;  2000US-0244617.
PR   08-NOV-2000;  2000US-0246474.

PR   08-NOV-2000;  2000US-0246475.
PR   08-NOV-2000;  2000US-0246476.
PR   08-NOV-2000;  2000US-0246477.
PR   08-NOV-2000;  2000US-0246478.
PR   08-NOV-2000;  2000US-0246523.
PR   08-NOV-2000;  2000US-0246524.
PR   08-NOV-2000;  2000US-0246525.
PR   08-NOV-2000;  2000US-0246526.
PR   08-NOV-2000;  2000US-0246527.
PR   08-NOV-2000;  2000US-0246528.
PR   08-NOV-2000;  2000US-0246532.
PR   08-NOV-2000;  2000US-0246609.
PR   08-NOV-2000;  2000US-0246610.
PR   08-NOV-2000;  2000US-0246611.
PR   08-NOV-2000;  2000US-0246613.
PR   17-NOV-2000;  2000US-0249207.
PR   17-NOV-2000;  2000US-0249208.
PR   17-NOV-2000;  2000US-0249209.
PR   17-NOV-2000;  2000US-0249210.
PR   17-NOV-2000;  2000US-0249211.
PR   17-NOV-2000;  2000US-0249212.
PR   17-NOV-2000;  2000US-0249213.
PR   17-NOV-2000;  2000US-0249214.
PR   17-NOV-2000;  2000US-0249215.
PR   17-NOV-2000;  2000US-0249216.
PR   17-NOV-2000;  2000US-0249217.
PR   17-NOV-2000;  2000US-0249218.
PR   17-NOV-2000;  2000US-0249244.
PR   17-NOV-2000;  2000US-0249245.
PR   17-NOV-2000;  2000US-0249264.
PR   17-NOV-2000;  2000US-0249265.
PR   17-NOV-2000;  2000US-0249297.
PR   17-NOV-2000;  2000US-0249299.
PR   17-NOV-2000;  2000US-0249300.
PR   17-NOV-2000;  2000US-0249300.
PR   01-DEC-2000;  2000US-0250160.
PR   01-DEC-2000;  2000US-0250391.
PR   05-DEC-2000;  2000US-0251030.
PR   05-DEC-2000;  2000US-0251988.
PR   05-DEC-2000;  2000US-0256719.
PR   06-DEC-2000;  2000US-0251479.
PR   08-DEC-2000;  2000US-0251856.
PR   08-DEC-2000;  2000US-0251868.
PR   08-DEC-2000;  2000US-0251869.
PR   08-DEC-2000;  2000US-0251989.
PR   08-DEC-2000;  2000US-0251990.
PR   11-DEC-2000;  2000US-0254097.
PR   05-JAN-2001;  2001US-0259678.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA, Barash SC, Ruben SM;
XX
DR   WPI; 2001-483426/52.
XX
PT   Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT   useful for preventing, diagnosing and/or treating cancers and
PT   metastasis -
XX
PS   Disclosure; SEQ ID NO 26580; 3071pp + Sequence Listing; English.
XX
CC   AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC   amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC   activity, and can be used in gene therapy and vaccine production. (I)
CC   proteins and polynucleotides may be used in the prevention, diagnosis and
CC   treatment of diseases associated with inappropriate (I) expression. For
CC   example, they may be used to treat disorders associated with decreased
CC   expression by rectifying mutations or deletions in a patient's genome
CC   that affect the activity of (I) by expressing inactive proteins or to
CC   supplement the patients own production of (I). Additionally, (I)
CC   polynucleotides may be used to produce the secreted (I), by inserting
CC   the nucleic acids into a host cell and culturing the cell to express the
CC   protein. (I) proteins and polynucleotides may be used to prevent,
CC   diagnose and treat immune/haematopoietic-related diseases, especially
```

```
CC    cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC    to AAK87694 represent human immune/haematopoietic antigen genomic
CC    sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC    represent sequences used in the exemplification of the present invention.
XX
SQ    Sequence 5262 BP; 1099 A; 1347 C; 1247 G; 1569 T; 0 other;

Query Match             35.5%;  Score 511.6;  DB 22;  Length 5262;
Best Local Similarity   70.3%;  Pred. No. 3.2e-126;
Matches  922;  Conservative   0;  Mismatches 299;  Indels  91;  Gaps  14;

Qy          1 tttttttttttgaagatgaagttttcgtttcttgtccaaggctggagtgcaatggcgtcaa 60
              |||||| ||||||||  ||||||||||||||||||||||||||||| || ||||||||
Db       1381 tttatgttttgagacagag-tctctctttgtgccaggctggaatgcaatggcgtga 1439

Qy         61 tctcagatcaccgcaacctccgcctccgggttcccaggttcaagggattcctccttcagcctcc 120
              ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db       1440 tctccggtcaactgcgcctccgggttcaagcgattcctcatgcctcagcctcc 1499

Qy        121 cagtagctcgggattacagcatgtcaccacgcctcggctaattttgtattttttag 180
              ||| ||||||||||||||||||||||||||||||||||||||||| | ||||||
Db       1500 cagtaggctaggattacaggcgtgcacgtgcaccatgttccagtc---tgtattttg 1556

Qy        181 tagaagatggagttttctccattgttggtcacgttctcagaactcccgacctcagatgatc 240
              ||| ||||  |||| ||||||||||||||||||||||||||||||||||||||||||||||
Db       1557 tagagacgggtgttcaccatgttgaccaggctgttctcaaactcctggga--tgttaccctg 1616

Qy        241 cctccgctctcggcctccacaaagtgctagatacagaggactggccaccatgcccggctctgcc 300
              ||| |||||||||||||||||||||||||||| || |||||| ||||| |||||||||
Db       1617 tgcccgcctggcctctcccaaaag-tgctgggattacaggcgtgaactacgcgcctggc 1675

Qy        301 tggctaatttttgtggtagaaaacaggttcactgatgtgtcccaagctggtctcctgagc 360
              |||| |  |||||||||||||||||||||||||||||||||| ||  ||||| ||| ||
Db       1676 -tttctcttttttcttttaatagacatcctgttgttaattctgttatggatttctt------ 1728

Qy        361 tcaagcagtccaacctgcctcagcctcccaaagtgctgggattacaggcgtcaagccgtgc 420
              ||||| |||||||||||||| ||||||||||||||||||||||||||||||----- 1728
Db       1729 gacagtggcaatgatcgcagctcactgcaacctccaagttcaagtcgattcct 1787

Qy        421 ctggccttttatttattttttaagacacaggttgtcccactcttaccaggatgaag 480
              ||| |||||||||||||||||||||||||||||||||| |||||||||||||||||
Db       1729 -tttctctttttttttattttttttaagacacaggttgtcccactcttaccaggatgaag 1728

Qy        481 tgcagtggtgtgatacaacagagtcgacttcaacctctgagatcaag-catcctcct 539
              | |||| |||||||||||||||||||||||||||||||||||||||||||||| ||||
Db       1788 gacagtggcaatgatcgcagctcactgcaacctccaagttcaagtcgattcct 1847

Qy        540 gcctccagctccccaagtagctgggaccaaagacatgcacacctacacctggctaattttgtatt 599
              ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db       1848 gcctccagctccccagtagctgggattacaggcatgcacaccacaccagctaattttt 1907

Qy        600 atttctattttttaattcttttgagacagagtcaactctgtcaacgggctggagtgcag 659
              ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db       1908 -tgtttgttttgttttgtttgatagggagttc-actctgtcacccaggctggagtcag 1965

Qy        660 tggcgcaatcttggctcactgcaacctctgcctcccgggttcaagtgattattctcctgccc 719
              |||| ||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db       1966 tcgtgcatcgttggctcactgcaacctccgcctcccaagtgctctcctctc 2025

Qy        720 agcctccgagtagctgggactacaggcgcccacacacgcctagctaat-tttttgtatt 778
              ||| |||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db       2026 agcctccgagtacctgggattaggacatgcgccacacatcagctaatttttgtatt 2085

Qy        779 ttagtaagatgggg--ttcacacatgttcgccagttgatcttgacct--tg 835
              ||||||||||||    ||||||||||||||||||||||||| |||||  ||
Db       2086 ttagtatatgacgggttcacacatgttggccagctggtcttgaactctgacctgagc 2145

Qy        836 tgatcctgcctgctcctcgcccttcccaaagtgctgggattacaggcgtgagccacacccg 895
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2146 tgatccgcctgctccccaaagtgctgggattacaggcatgagccaccgctgcctc 2205

Qy        896 gcttattttaattttgtttgtttgaaatggaatctcactc---tgttaccaggctgtg 952
              ||||||||||||||||||||||||||||||||||||||||   |||||||||||||||||
```

```
Db       2206 gcctgtt------tgtttcttttttggagtggtctcgctcttgttccgccaggctgg 2260
              ||  |||      |||||| |||||||||||||||||||||||| |||||| |||||| |||
Qy       1012 agtgcaatggcgcaatctcggcctcactgcaacctctgcctccgggttcaagcgattctc 1012

Db       2261 agtgcagtggcgcgatctcggctcactgcaacctccactt--tgggttcaagcgattccact 2318
              ||||||| |||||| |||||||||||||||||||||||| |   |||||||||||||||||
Qy       2318 ctgcctcagcctcccaagtagctgggattacaggcgccaccaccacccgcctaatttt 2318

Db       2319 ctgcctcagcct-cgaagtagctgggattacaggcgtgcaccaccatcggttaaatt 2377
              |||||||||||| ||||||||||||||||||||||||||| ||||| || |||| |
Qy       2377 ttgtattttttagtagagacggggtttcaccatattggtcaggctggtctcgaactcctga 2377

Db       1073 ttgtattttcattagagacgcggggttcaccacatattgtcaggctggtctccaaactcctga 1132
              |||||||| |||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Qy       1133 cctcaggtgaccacctgcctcagcctcccaaagtgctgggattacaggcgtgagccac 1192

Db       2378 ttgtaattttagtagagacgggtttcaccatgttggccaggctggtctcgaactcctga 2437
              |||| ||||||||||||||||||| |||||||||| ||||||||||||||||||||||||
Qy       1192 cctcaggtgatccgcccgcctcggcctcccaaagtgctgggattacaggcgtgagccact 2497

Db       2438 cctcaagtgatccgcccgccttcggcctcccaaagtgctgggattacaggcgtgagccact 2497
              ||||| |||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Qy       1252 tcaccacgcgggctaattttagataaaaatatgtacggagcaatgtggggggtcttgctatgttg 1252

Db       1193 tcaccacgcgggctaatttttagataaaaatatgtacggagcaatgtggggggtcttgctatgttg 1252
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       1304 cccaggctggtctccaaactctggcttcatgcaatcctcccaaatgagcac 1304

Db       2498 gctcctggccccactggccgctaatttttgtatttttagtagagatggggtttcaccatgttg 2557
              ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Qy       2609 cccaggctggtctccaaactcctggcttcatgcaatcctcccaaatgagcac 2609

Db       2558 gccaggctggtctgcaagtgttatccgcctcggcctcccaaagtgctgagcctc 2609
```

```
RESULT   7
AAK71769
ID   AAK71769    standard; DNA; 5262 BP.
XX
AC   AAK71769;
XX
DT   06-NOV-2001   (first entry)
XX
DE   Human immune/haematopoietic antigen genomic sequence SEQ ID NO:26581.
XX
KW   Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW   cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS   Homo sapiens.
XX
XX
PN   WO200157182-A2.
XX
PD   09-AUG-2001.
XX
PF   17-JAN-2001;  2001WO-US01354.
XX
PR   31-JAN-2000;  2000US-0179065.
PR   04-FEB-2000;  2000US-0180628.
PR   24-FEB-2000;  2000US-0184664.
PR   02-MAR-2000;  2000US-0186350.
PR   16-MAR-2000;  2000US-0189874.
PR   17-MAR-2000;  2000US-0190076.
PR   18-APR-2000;  2000US-0198123.
PR   19-MAY-2000;  2000US-0205515.
PR   07-JUN-2000;  2000US-0209467.
PR   28-JUN-2000;  2000US-0214886.
PR   30-JUN-2000;  2000US-0215135.
PR   07-JUL-2000;  2000US-0216647.
PR   07-JUL-2000;  2000US-0216880.
PR   11-JUL-2000;  2000US-0217487.
PR   11-JUL-2000;  2000US-0217496.
PR   14-JUL-2000;  2000US-0218290.
PR   26-JUL-2000;  2000US-0220963.
PR   26-JUL-2000;  2000US-0220964.
PR   14-AUG-2000;  2000US-0224518.
PR   14-AUG-2000;  2000US-0224519.
PR   14-AUG-2000;  2000US-0225213.
PR   14-AUG-2000;  2000US-0225214.
PR   14-AUG-2000;  2000US-0225266.
PR   14-AUG-2000;  2000US-0225267.
```

```
PR   14-AUG-2000;  2000US-0225268.
PR   14-AUG-2000;  2000US-0252270.
PR   14-AUG-2000;  2000US-0225447.
PR   14-AUG-2000;  2000US-0225757.
PR   14-AUG-2000;  2000US-0255758.
PR   14-AUG-2000;  2000US-0255759.
PR   18-AUG-2000;  2000US-0226279.
PR   22-AUG-2000;  2000US-0226681.
PR   22-AUG-2000;  2000US-0226868.
PR   22-AUG-2000;  2000US-0227182.
PR   23-AUG-2000;  2000US-0227009.
PR   30-AUG-2000;  2000US-0228924.
PR   01-SEP-2000;  2000US-0229287.
PR   01-SEP-2000;  2000US-0229343.
PR   01-SEP-2000;  2000US-0229344.
PR   01-SEP-2000;  2000US-0229345.
PR   05-SEP-2000;  2000US-0229509.
PR   05-SEP-2000;  2000US-0229513.
PR   06-SEP-2000;  2000US-0230437.
PR   06-SEP-2000;  2000US-0230438.
PR   08-SEP-2000;  2000US-0231242.
PR   08-SEP-2000;  2000US-0231243.
PR   08-SEP-2000;  2000US-0231244.
PR   08-SEP-2000;  2000US-0231413.
PR   08-SEP-2000;  2000US-0231414.
PR   08-SEP-2000;  2000US-0232081.
PR   12-SEP-2000;  2000US-0231968.
PR   14-SEP-2000;  2000US-0232297.
PR   14-SEP-2000;  2000US-0232298.
PR   14-SEP-2000;  2000US-0232299.
PR   14-SEP-2000;  2000US-0232400.
PR   14-SEP-2000;  2000US-0232401.
PR   14-SEP-2000;  2000US-0233063.
PR   14-SEP-2000;  2000US-0233064.
PR   14-SEP-2000;  2000US-0233065.
PR   21-SEP-2000;  2000US-0234223.
PR   21-SEP-2000;  2000US-0234274.
PR   25-SEP-2000;  2000US-0234997.
PR   25-SEP-2000;  2000US-0234998.
PR   26-SEP-2000;  2000US-0235484.
PR   27-SEP-2000;  2000US-0235834.
PR   27-SEP-2000;  2000US-0235836.
PR   29-SEP-2000;  2000US-0236327.
PR   29-SEP-2000;  2000US-0236367.
PR   29-SEP-2000;  2000US-0236368.
PR   29-SEP-2000;  2000US-0236369.
PR   29-SEP-2000;  2000US-0236370.
PR   02-OCT-2000;  2000US-0236802.
PR   02-OCT-2000;  2000US-0237037.
PR   02-OCT-2000;  2000US-0237038.
PR   02-OCT-2000;  2000US-0237039.
PR   02-OCT-2000;  2000US-0237040.
PR   13-OCT-2000;  2000US-0239935.
PR   13-OCT-2000;  2000US-0239937.
PR   20-OCT-2000;  2000US-0240960.
PR   20-OCT-2000;  2000US-0241221.
PR   20-OCT-2000;  2000US-0241785.
PR   20-OCT-2000;  2000US-0241786.
PR   20-OCT-2000;  2000US-0241787.
PR   20-OCT-2000;  2000US-0241808.
PR   20-OCT-2000;  2000US-0241809.
PR   20-OCT-2000;  2000US-0241826.
PR   01-NOV-2000;  2000US-0244417.
PR   08-NOV-2000;  2000US-0244474.
PR   08-NOV-2000;  2000US-0244475.
PR   08-NOV-2000;  2000US-0244476.
PR   08-NOV-2000;  2000US-0244477.
PR   08-NOV-2000;  2000US-0244478.
PR   08-NOV-2000;  2000US-0246476.
PR   08-NOV-2000;  2000US-0246477.
PR   08-NOV-2000;  2000US-0246478.
PR   08-NOV-2000;  2000US-0246523.
PR   08-NOV-2000;  2000US-0246524.
PR   08-NOV-2000;  2000US-0246525.
PR   08-NOV-2000;  2000US-0246526.
PR   08-NOV-2000;  2000US-0246527.
PR   08-NOV-2000;  2000US-0246528.
PR   08-NOV-2000;  2000US-0246532.
PR   08-NOV-2000;  2000US-0246609.
PR   08-NOV-2000;  2000US-0246610.
PR   08-NOV-2000;  2000US-0246611.
PR   09-NOV-2000;  2000US-0246613.
PR   17-NOV-2000;  2000US-0249207.
PR   17-NOV-2000;  2000US-0249208.
PR   17-NOV-2000;  2000US-0249209.
PR   17-NOV-2000;  2000US-0249210.
PR   17-NOV-2000;  2000US-0249211.
PR   17-NOV-2000;  2000US-0249212.
PR   17-NOV-2000;  2000US-0249213.
PR   17-NOV-2000;  2000US-0249214.
PR   17-NOV-2000;  2000US-0249215.
PR   17-NOV-2000;  2000US-0249216.
PR   17-NOV-2000;  2000US-0249217.
PR   17-NOV-2000;  2000US-0249218.
PR   17-NOV-2000;  2000US-0249244.
PR   17-NOV-2000;  2000US-0249245.
PR   17-NOV-2000;  2000US-0249264.
PR   01-DEC-2000;  2000US-0249265.
PR   01-DEC-2000;  2000US-0250160.
PR   01-DEC-2000;  2000US-0250391.
PR   05-DEC-2000;  2000US-0251030.
PR   05-DEC-2000;  2000US-0251988.
PR   05-DEC-2000;  2000US-0256719.
PR   06-DEC-2000;  2000US-0251479.
PR   08-DEC-2000;  2000US-0251856.
PR   08-DEC-2000;  2000US-0251868.
PR   08-DEC-2000;  2000US-0251869.
PR   08-DEC-2000;  2000US-0251989.
PR   08-DEC-2000;  2000US-0251990.
PR   11-DEC-2000;  2000US-0254097.
PR   05-JAN-2001;  2001US-0255678.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA, Barash SC, Ruben SM;
XX
DR   WPI; 2001-483426/52.
XX
PT   Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT   useful for preventing, diagnosing and/or treating cancers and
PT   metastasis -
XX
PS   Disclosure; SEQ ID NO 26581; 3071pp + Sequence Listing; English.
XX
CC   AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC   amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC   activity, and can be used in gene therapy and vaccine production. (I)
CC   proteins and polynucleotides may be used in the prevention, diagnosis and
CC   treatment of diseases associated with inappropriate (I) expression. For
CC   example, they may be used to treat disorders associated with decreased
CC   expression by rectifying mutations or deletions in a patient's genome
CC   that affect the activity of (I) by expressing inactive proteins or to
CC   supplement the patients own production of (I). Additionally, (I)
CC   polynucleotides may be used to produce the secreted (I), by inserting
CC   the nucleic acids into a host cell and culturing the cell to express the
CC   protein. (I) proteins and polynucleotides may be used to prevent,
CC   diagnose and treat immune/haematopoietic-related diseases, especially
CC   cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC   to AAK87694 represent human immune/haematopoietic antigen genomic
CC   sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC   represent sequences used in the exemplification of the present invention.
XX
SQ   Sequence 5262 BP; 1100 A; 1346 C; 1245 G; 1571 T; 0 other;
```

```
Query Match    35.5%;  Score 511.6;  DB 22;  Length 5262;
Best Local Similarity  70.3%;
Matches  922;  Conservative   0;  Mismatches 299;  Indels  91;  Gaps  14;

Qy        1 tttttttttgagatggagtttcgcttcttgtcgcccgggctggagtggcgcaa   60
            ||||||||||||||||||||| ||| ||||||||||||| |||||||||||||
Db     1381 ttttatgttttgagacagag-tctctctcttgttcgcccaggctggaatgcagtggcgtga 1439

Qy       61 tctcagctcaccgcaacctcgcctcccgggttcaagcgattctcctgcctcagcctcc   120
            |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1440 tctcggctcactgcaacctccgcctcccaggttcaagcgattctcatgcctcagcctccc 1499

Qy      121 cagtagctgggattacaggcgcatgccaccacgctcggctaattttttgtattttttag   180
            |||||||||||||||||||| |||||||||| ||||||||||||||||||||||||
Db     1500 cagtagctaggattacaggcgcaccacgcccagctatttt---tgtattttgg 1556

Qy      181 tagagatggagtttctcatgttggttcaggctgtctcgaacctccgacctcagatgatc   240
            |||||||||| |||||||||||||| |||||||||||||||||||| ||||||
Db     1557 tagagacggggtttcaccatgttgacaggctggtctcgaactcctgacctcagtgatc 1616

Qy      241 cctccgtctcggcctcccaaagtgctagatacagactggccaccatgcccggctctgcc   300
            ||||||||||||||||||||||| | |||||||||| ||||||||||||||||
Db     1617 tgcccgccttggcctctcccaaag-tgctgggattacaggcgtgaactacctggc 1675

Qy      301 tggctaattttgtgtggtagaaacaaggttcactgatgtgccaaggttctcctgagc   360
            ||||||| |||||||||||| |||||||||||||||||||||||| ||||||
Db     1676 ctgcatgtttctctaataagcatcctgtgttaattctgtatggacttctt------ 1728

Qy      361 tcaagcagtcacctgcctcagcctccaaagtgtgggattacaggcgtgcagccgtgc   420
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1729 -------------------------------------------------------- 1728

Qy      421 ctggcctcttttatttattttttaagacacaagtgtccactcttaccaggatgaag   480
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1729 -tttcttcttttattttttgagacagagtctcgccttgtcacccaggctggag 1787

Qy      481 tgcagtggtgtggatcacaagctcactgcagccttcaaagctgagatcaag-catcctcct 539
            |||| ||||||||||||||||||||||||||||||||||| |||||||||
Db     1788 gacagtggcatgatcggctcactgcaacctctgcctcccaattccaggttcaagtgattctcct 1847

Qy      540 gcctcagcctagctgggaccaaagacatgcaccactacaccctggctaatttt   599
            ||||||||||||||||||||| |||||||||||||||||||| |||||||
Db     1848 gcctcagcctcccaagtagctgggattacaggcatgcaccaccacccagctaatttt 1907

Qy      600 atttttattttttaattctttgagacagagtctcaactctgtcaactggagtgcag   659
            |||||||||||||||| |||||||||||| |||||||||||||||
Db     1908 -tgttttgttttgttttttgatagggagttc-actgtgcacccaggctggagttgcag 1965

Qy      660 tggcgcaatcttggctcactgcaacctctgcctcccgggttcaagtgattattctccgcc   719
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1966 tcgtgccatgttggctcactgcaacctccgcctcccaggttcaagtgattctcctgtctc 2025

Qy      720 agcctcctgagtgctgggactacaggcgcccaccacgcctagctaat-ttttgtatt   778
            |||||||||||||||||||||||||||||||||||||||| ||||||||||
Db     2026 agcctcgagtaacctgggattataggcgccccatgccagctcctaatttttgtatt 2085

Qy      779 tttagtagagaatggggg-ttcaccatgttccccaggttgatctctgacct--tg   835
            ||||||||||||||||| |||||||||| |||||||||||||||||||
Db     2086 tttagtagacacggggtttcaccatgttggccaggctggtcttgaacctcctgacctcaag 2145

Qy      836 tgatctgtgcctcggcctcccaaagtgctgggattatgagccaccacgcccg   895
            |||||||||||||||||||||||||| |||||||||||||||||||||||
Db     2146 tgatccgcctcagcctcccaaagtgctgggattacaggcgtgagccaccacgcccg 2205

Qy      896 gcttattttttaaatttgttttgaaatcacactc---tgttaccaggctgg   952
            ||||||||| |||||||||||||||||||  || ||||||||||||||
Db     2206 gcctgtt----tgttcttttttgagaagagtctcggctctgttgcccaggctgg 2260

Qy      953 agtgcaatgccaaatctcggctcactgcctccgggttcaagcgattctc  1012
            ||||| ||| ||||||||||||||||||||||||||||| |||||
Db     2261 agtgcagtgacgatccggctcactgcaactcactt--tgggttcaagcgattctc 2318

Qy     1013 ctgtctcagcctcccaagcagctgggattacgggcacctgccaccaccccgcctaatt  1072
            ||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db     2319 ctgcctcagcct-cgaagtagctgggattacaggcatgccacacactcggttaaatt 2377

Qy     1073 ttgtatttcattagaggcggggttttcaccatatttgtcaggctggt  1132
            |||||||||||||||||||||||||||||||||||||||||||||
Db     2378 ttgtatttttagtagagacaggggttatcggatgtgtctcaaactcctga 2437

Qy     1133 cctcaggtgaccaccctgcctcagcctcccaaagtgctgggattacagtgat  1192
            |||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db     2438 cctcaagtgatccgcccgcctcggcctcccaaagtgctgggattacaggcgtgagccact 2497

Qy     1193 tcaccacgccgctaatttagataaaaaatatgtacaatgggggtcttgttct  1252
            ||||||||| |||||||||||||||||||||||||||||||||||||||
Db     2498 gcctcctgggccaacctggtctaatttttgtatttttagtagagatggggtttcttcat 2557

Qy     1253 cccaggctgggtctcaaacttctgcttcatgtcaatcctccaaatgagccac  1304
            ||| |||||||||||||||||||||||||||||||||| |||||||||||||
Db     2558 gccaggctgttctcgaagtcctgacctcaggtgatccacctgcctcagcctc 2609
```

RESULT .8

```
PR   14-AUG-2000;  2000US-0225757.
PR   14-AUG-2000;  2000US-0225758.
PR   14-AUG-2000;  2000US-0225759.
PR   18-AUG-2000;  2000US-0226279.
PR   22-AUG-2000;  2000US-0226681.
PR   22-AUG-2000;  2000US-0226868.
PR   22-AUG-2000;  2000US-0227182.
PR   23-AUG-2000;  2000US-0227009.
PR   30-AUG-2000;  2000US-0228924.
PR   01-SEP-2000;  2000US-0229287.
PR   01-SEP-2000;  2000US-0229343.
PR   01-SEP-2000;  2000US-0229344.
PR   01-SEP-2000;  2000US-0229345.
PR   05-SEP-2000;  2000US-0229509.
PR   05-SEP-2000;  2000US-0229513.
PR   06-SEP-2000;  2000US-0230437.
PR   06-SEP-2000;  2000US-0230438.
PR   08-SEP-2000;  2000US-0231242.
PR   08-SEP-2000;  2000US-0231243.
PR   08-SEP-2000;  2000US-0231244.
PR   08-SEP-2000;  2000US-0231413.
PR   08-SEP-2000;  2000US-0231414.
PR   08-SEP-2000;  2000US-0232080.
PR   08-SEP-2000;  2000US-0232081.
PR   12-SEP-2000;  2000US-0231968.
PR   14-SEP-2000;  2000US-0232397.
PR   14-SEP-2000;  2000US-0232398.
PR   14-SEP-2000;  2000US-0232399.
PR   14-SEP-2000;  2000US-0232400.
PR   14-SEP-2000;  2000US-0232401.
PR   14-SEP-2000;  2000US-0233063.
PR   14-SEP-2000;  2000US-0233064.
PR   14-SEP-2000;  2000US-0233065.
PR   21-SEP-2000;  2000US-0234223.
PR   21-SEP-2000;  2000US-0234274.
PR   25-SEP-2000;  2000US-0234997.
PR   25-SEP-2000;  2000US-0234998.
PR   26-SEP-2000;  2000US-0235484.
PR   27-SEP-2000;  2000US-0235834.
PR   27-SEP-2000;  2000US-0235836.
PR   29-SEP-2000;  2000US-0236327.
PR   29-SEP-2000;  2000US-0236367.
PR   29-SEP-2000;  2000US-0236368.
PR   29-SEP-2000;  2000US-0236369.
PR   29-SEP-2000;  2000US-0236370.
PR   02-OCT-2000;  2000US-0236802.
PR   02-OCT-2000;  2000US-0237037.
PR   02-OCT-2000;  2000US-0237038.
PR   02-OCT-2000;  2000US-0237039.
PR   02-OCT-2000;  2000US-0237040.
PR   13-OCT-2000;  2000US-0239935.
PR   13-OCT-2000;  2000US-0239937.
PR   20-OCT-2000;  2000US-0240960.
PR   20-OCT-2000;  2000US-0241785.
PR   20-OCT-2000;  2000US-0241786.
PR   20-OCT-2000;  2000US-0241787.
PR   20-OCT-2000;  2000US-0241808.
PR   20-OCT-2000;  2000US-0241809.
PR   20-OCT-2000;  2000US-0241826.
PR   20-OCT-2000;  2000US-0242221.
PR   01-NOV-2000;  2000US-0246474.
PR   08-NOV-2000;  2000US-0246474.
PR   08-NOV-2000;  2000US-0246475.
PR   08-NOV-2000;  2000US-0246476.
PR   08-NOV-2000;  2000US-0246477.
PR   08-NOV-2000;  2000US-0246478.
PR   08-NOV-2000;  2000US-0246523.
PR   08-NOV-2000;  2000US-0246524.
PR   08-NOV-2000;  2000US-0246525.
PR   08-NOV-2000;  2000US-0246526.
PR   08-NOV-2000;  2000US-0246527.
PR   08-NOV-2000;  2000US-0246528.
PR   08-NOV-2000;  2000US-0246532.
PR   08-NOV-2000;  2000US-0246609.
PR   08-NOV-2000;  2000US-0246610.
PR   08-NOV-2000;  2000US-0246611.
PR   08-NOV-2000;  2000US-0246613.
PR   17-NOV-2000;  2000US-0249207.
PR   17-NOV-2000;  2000US-0249208.
PR   17-NOV-2000;  2000US-0249209.
PR   17-NOV-2000;  2000US-0249210.
PR   17-NOV-2000;  2000US-0249211.
PR   17-NOV-2000;  2000US-0249212.
PR   17-NOV-2000;  2000US-0249213.
PR   17-NOV-2000;  2000US-0249214.
PR   17-NOV-2000;  2000US-0249215.
PR   17-NOV-2000;  2000US-0249216.
PR   17-NOV-2000;  2000US-0249217.
PR   17-NOV-2000;  2000US-0249218.
PR   17-NOV-2000;  2000US-0249244.
PR   17-NOV-2000;  2000US-0249245.
PR   17-NOV-2000;  2000US-0249264.
PR   17-NOV-2000;  2000US-0249265.
PR   17-NOV-2000;  2000US-0249297.
PR   17-NOV-2000;  2000US-0249299.
PR   17-NOV-2000;  2000US-0249300.
PR   01-DEC-2000;  2000US-0250391.
PR   01-DEC-2000;  2000US-0251160.
PR   05-DEC-2000;  2000US-0251030.
PR   05-DEC-2000;  2000US-0251988.
PR   05-DEC-2000;  2000US-0256719.
PR   06-DEC-2000;  2000US-0251479.
PR   08-DEC-2000;  2000US-0251856.
PR   08-DEC-2000;  2000US-0251868.
PR   08-DEC-2000;  2000US-0251869.
PR   08-DEC-2000;  2000US-0251989.
PR   08-DEC-2000;  2000US-0251990.
PR   11-DEC-2000;  2000US-0254097.
PR   05-JAN-2001;  2001US-0259678.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA, Barash SC, Ruben SM;
XX
DR   WPI; 2001-541565/60.
XX
PT   Nucleic acids encoding 3224 human nervous system antigen polypeptides,
PT   useful for preventing, diagnosing and/or treating nervous system
PT   cancers and metastases -
XX
PS   Disclosure; SEQ ID NO 8463; 1701pp + Sequence Listing; English.
XX
CC   The invention relates to novel genes (ABA11004-ABA21534) and proteins
CC   (ABB14678-ABB18001) useful for preventing, treating or ameliorating
CC   medical conditions e.g. by protein or gene therapy. The genes are
CC   isolated from a range of human tissues disclosed in the specification.
CC   The nucleic acids, proteins, antibodies and (ant)agonists are useful
CC   in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC   and ovarian cancer and other cancers of the adrenal gland, bone,
CC   marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC   (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC   haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC   disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC   colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC   (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC   epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC   and parasitic infections.
CC   Note: The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 24167 BP; 6223 A; 6198 C; 6379 G; 5367 T; 0 other;
XX

Query Match          34.0%;   Score 489.6;   DB 22;   Length 24167;
Best Local Similarity 69.8%;   Pred. No. 4.3e-120;
```

Matches 814;   Conservative 0;   Mismatches 319;   Indels 33;   Gaps 10;

```
Qy      62 ctcaagctcaacgcaaactcccgcctcccgggttccaaagacgattctctgcctcaagcctcccc   121
           ||||| |  | |  |||||||||  ||||| || ||||||||||| || | ||||||| |||||||
Db   14146 CTTGGGCTCACACTGCAACTCCGCCTCTTGGGTTCAAGCGCTTCTTCTGCCCTCAGCCTCCCC   14087

Qy     122 agtagctgggattacaggcatgtgacccacacgctcggctcaatttgtattttttttagt       181
           | ||||||||||  |||| | |  |  ||||| || || |||||| || ||||||||||  |
Db   14086 AGTAGCTGGGACTATAGGCACGTCCCACTACGCCCACCGCTCAATTTT TTGTATTTTTTAGT   14028

Qy     182 agagatggagttctccatgttggtcaggctgtctcgaactcccgacctcaagtgatcc       241
           |||| | ||| |||| | ||  |||| | || | ||   ||||| ||| ||| || |||
Db   14027 AGAGATGGGGTTTCACTATAGTTAGCCAGGATGGTCTCGAACTCCTGACCTCGTGATCTGC   13968

Qy     242 ctccgtctcggcctcccaaagtgctagatacaggcgtgccacatgccc--ggctctgc     299
           || ||| || | |||||||||||| | || ||||| |  |||| || ||  || ||||
Db   13967 CTGCCTTGGGCTCTCCAAAGTGTTGGGATTATAGGCACCACTGCCCTTTTT TGGCCTTTTT   13908

Qy     300 ctggctcaattttgtggtagaaacaggtttcactgatgtgccaaagctggtctcctgag     359
           || |||||| || || |  |  |||| | || | |||||||| ||||||||||| ||
Db   13907 TTTTTTTTTTTTTTTTGAGATGGAGTTTTCGCTCTTGTTGCCCAGGCTGGAGTGCAATGGTG   13848

Qy     360 ctcaagcagtcaacctgcctcccaaagtgctgggattacaggcgtgcagccgtg     419
           || |||| || | |||| |||||||||||||||||||||||| |||| |  |||
Db   13847 ACTCGGCTCACTGCAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCC   13788

Qy     420 cctggcctttttatttttttttttaagacacaggtgtccacttttaccagggatgaa     479
           |  ||| |||||||  ||||||||||   |||| | |||| |||  || ||| |
Db   13787 CGAGTAGCTGGGATTACAGGCGCACCACCACGCCCGGCTAATTTTTGTACTTTTTAGTA   13728

Qy     480 gtgtcagttgtgtgatca--cagtcactgcagccttcaactcctgagatcaag-catct     536
           | | | ||| |  ||||    || | | | ||  | || ||| ||| ||||| |  |||
Db   13727 GAGACGGGGTTTCTCCATGTTGGTCAGGCTGGTCTCAAACTCCTGACCTCAGGTGATCCA   13668

Qy     537 cctgcctcagcctcccaagtagctgggaccaaagacagtcacactacacctgc-----     591
           ||||||| ||||||||||||  ||| |||  ||  | | | | ||| ||| |||
Db   13667 CCCACCTCGGCCTCCCAAAGTGTTGGGATTACAGACATGAGCCACCGCCTGGCCGTGT   13608

Qy     592 --------------taattttttattttttttagagacagagtctcaact        636
                         | |||||||| ||| |||||||||| |||| | ||
Db   13607 CTGGCCTTTTTTAGTATTTCTTTTTTTTTTTTTTTTGAGACAGAGTCTT-TACT   13549

Qy     637 ctgtcaccaagctggagtgcagtggcgccaatctggctcactgcaacctctgcctccg     696
           || | ||||||||||||||||| ||| | || || || |||||||  ||||| | |||
Db   13548 CCGTCGCCCAGGCTGGAGTGCAGTGGCTGCGATGTCTGCGGCTCACTGCAGCCTCCGCCTCCCG   13489

Qy     697 ggttcaagttattcctcctgccccagcctcccgagtagctgggactacaggcgcccaccac     756
           |||||||| ||||| ||| |||| ||||| ||| |||| | |||||||||| |  ||| ||
Db   13488 GGTTCAAGTCAATTCTCCTGCCTCAGCCTCCGAGTAGCTGGGACTACAGGCGCCCGCCAC   13429

Qy     757 --gcctagctaattttttgtattttttagtagagacgatgg-ggttcaccatgttgccaagt     813
             |||| | |||||||||| | ||||||||| ||| |  |  ||| |||||| ||| ||||
Db   13428 TACGCCCGGCTACTTTTTGTATATTTAGTAGAGATGGGGTTTCACCATGTTAGCCAGGA   13369

Qy     814 tgatcttgatctggaccttgtctcggcgtctcggcgtctcggggattca        873
           ||||| ||| ||| ||||| || |  | ||||  || || | |||| |
Db   13368 TGGTCTCGATCTCCTGACCTTGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTA   13309

Qy     874 caggcgcgtgagccaccacgcccggccttatttttaatttttgtttgaaatggaatctc     933
           |||||| ||||| ||||| |  | |||  |||||||| ||| || |  ||  ||| |||
Db   13308 CAGGCGTGAGCCACCACATGCCCAGGCTT-TTTTTTTTTTTTGTTTTTTGAGACGGAGTCTT   13250

Qy     934 actctgttacccaggctggagtgcaaatcggcctcactgcaacctctgcctc     993
           |||| ||| | | ||||||||||||| ||||| |||||||||  |||| | ||
Db   13249 GCTCTGTTGCCCAGGCTGGAGTGCAGTGGCGCCATGATCTCAGCTCACTGCAAGCTCCACTTC   13190

Qy     994 ccgggctcaagcgagattcctgtctcagcctcccgaacagctgggattacgggcacctgc     1053
           || ||||| |||| || | || |||||||||||| |||| || ||||| |  |  || |
Db   13189 CCAGGCTCAAGCGATTCTCCAGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGCGCCCGC   13130

Qy     1054 caccacacccgctaa--ttttgtattttcattagagacgggggtttcaccatatttgtc     1111
            | ||||||| ||||   |||| ||||| || |||||| || || |||||||| ||||
Db   13129 CACCACACTCGCTAATTTTTTGTATTTTTAGTAGAGACGGGGTTTCACCATGTTAGCC   13070

Qy     1112 aggctggtctccaaactcctgacctcaggtgacccacctgcctcagcctcccaaagtgctg     1171
            ||||||| |||||||||||||||||||  ||||| ||||| ||||||||||||||| ||
Db   13069 AGGCTGGTCTTGAACTCCTGACCTCAGGCCGATTCACCTGCCTCGGCCTCCCAAAGTGCTG   13010

Qy     1172 ggattacaggcgtgagccacctcacc     1197
           |||||||||| | |||| ||||| ||
Db   13009 GGATTAAAGGTGTGAGCCACCACCTCGCC   12984
```

PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
PR   08-SEP-2000; 2000US-0232080.
PR   08-SEP-2000; 2000US-0232081.
PR   12-SEP-2000; 2000US-0231968.
PR   14-SEP-2000; 2000US-0232297.
PR   14-SEP-2000; 2000US-0232298.
PR   14-SEP-2000; 2000US-0232399.
PR   14-SEP-2000; 2000US-0232400.
PR   14-SEP-2000; 2000US-0232401.
PR   14-SEP-2000; 2000US-0233063.
PR   14-SEP-2000; 2000US-0233064.
PR   14-SEP-2000; 2000US-0233065.
PR   21-SEP-2000; 2000US-0234223.
PR   21-SEP-2000; 2000US-0234274.
PR   25-SEP-2000; 2000US-0234997.
PR   25-SEP-2000; 2000US-0234998.
PR   26-SEP-2000; 2000US-0235484.
PR   27-SEP-2000; 2000US-0235834.
PR   27-SEP-2000; 2000US-0235836.
PR   29-SEP-2000; 2000US-0236327.
PR   29-SEP-2000; 2000US-0236367.
PR   29-SEP-2000; 2000US-0236368.
PR   29-SEP-2000; 2000US-0236369.
PR   29-SEP-2000; 2000US-0236370.
PR   13-OCT-2000; 2000US-0239935.
PR   02-OCT-2000; 2000US-0237040.
PR   02-OCT-2000; 2000US-0236802.
PR   02-OCT-2000; 2000US-0237037.
PR   02-OCT-2000; 2000US-0237038.
PR   02-OCT-2000; 2000US-0237039.
PR   20-OCT-2000; 2000US-0241221.
PR   20-OCT-2000; 2000US-0240960.
PR   20-OCT-2000; 2000US-0241809.
PR   20-OCT-2000; 2000US-0241826.
PR   01-OCT-2000; 2000US-0244617.
PR   20-OCT-2000; 2000US-0241785.
PR   20-OCT-2000; 2000US-0241786.
PR   20-OCT-2000; 2000US-0241787.
PR   20-OCT-2000; 2000US-0241808.
PR   08-NOV-2000; 2000US-0246474.
PR   08-NOV-2000; 2000US-0246475.
PR   08-NOV-2000; 2000US-0246476.
PR   08-NOV-2000; 2000US-0246477.
PR   08-NOV-2000; 2000US-0246478.
PR   08-NOV-2000; 2000US-0246523.
PR   08-NOV-2000; 2000US-0246524.
PR   08-NOV-2000; 2000US-0246525.
PR   08-NOV-2000; 2000US-0246526.
PR   08-NOV-2000; 2000US-0246527.
PR   08-NOV-2000; 2000US-0246528.
PR   08-NOV-2000; 2000US-0246532.
PR   08-NOV-2000; 2000US-0246609.
PR   08-NOV-2000; 2000US-0246610.
PR   08-NOV-2000; 2000US-0246611.
PR   08-NOV-2000; 2000US-0246613.
PR   17-NOV-2000; 2000US-0249207.
PR   17-NOV-2000; 2000US-0249208.
PR   17-NOV-2000; 2000US-0249209.
PR   17-NOV-2000; 2000US-0249210.
PR   17-NOV-2000; 2000US-0249211.

PR   17-NOV-2000; 2000US-0249212.
PR   17-NOV-2000; 2000US-0249213.
PR   17-NOV-2000; 2000US-0249214.
PR   17-NOV-2000; 2000US-0249215.
PR   17-NOV-2000; 2000US-0249216.
PR   17-NOV-2000; 2000US-0249217.
PR   17-NOV-2000; 2000US-0249218.
PR   17-NOV-2000; 2000US-0249244.
PR   17-NOV-2000; 2000US-0249245.
PR   17-NOV-2000; 2000US-0249264.
PR   17-NOV-2000; 2000US-0249265.
PR   17-NOV-2000; 2000US-0249297.
PR   17-NOV-2000; 2000US-0249299.
PR   17-NOV-2000; 2000US-0249300.
PR   01-DEC-2000; 2000US-0250160.
PR   01-DEC-2000; 2000US-0250391.
PR   05-DEC-2000; 2000US-0251030.
PR   05-DEC-2000; 2000US-0251988.
PR   05-DEC-2000; 2000US-0256719.
PR   06-DEC-2000; 2000US-0251479.
PR   08-DEC-2000; 2000US-0251856.
PR   08-DEC-2000; 2000US-0251868.
PR   08-DEC-2000; 2000US-0251869.
PR   08-DEC-2000; 2000US-0251989.
PR   08-DEC-2000; 2000US-0251990.
PR   11-DEC-2000; 2000US-0254097.
PR   05-JAN-2001; 2001US-0259678.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA, Barash SC, Ruben SM;
XX
DR   WPI; 2001-465569/50.
XX
PT   Isolated nucleic acid molecule encoding excretory system antigen is
     used in preventing, treating or ameliorating a medical condition -
XX
PS   Example 2; SEQ ID NO 757; 574pp + Sequence Listing; English.
XX
CC   The invention relates to novel excretory system related human
CC   polynucleotides (AAI98567-AAI99503) and the encoded proteins
CC   (AAM99594-AAM99913) useful for preventing, treating or ameliorating
CC   medical conditions e.g. by protein or gene therapy, especially
CC   disorders related to the excretory system. The genes are isolated
CC   from a range of human tissues disclosed in the specification. The
CC   nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC   the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC   and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC   marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC   (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC   haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC   disease, multiple sclerosis, rheumatoid arthritis, diabetes mellitus, ulcerative
CC   colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC   (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC   epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC   and parasitic infections.
CC   Note: The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 31853 BP; 8092 A; 6947 C; 7314 G; 9500 T; 0 other;

     Query Match   33.2%; Score 478.2; DB 22; Length 31853;
     Best Local Similarity   70.8%;
     Matches 890; Conservative   0; Mismatches 313; Indels   54; Gaps 17;
                              Pred. No. 5.3e-117;

Qy   1 tttttttttgagatggagttttcgctcttgttgcccaggctggagtgcaatgggcgcaa   60
             ||||||||||||||||||||||||||| |||| ||| ||||||||||||||||| ||||
Db   3208 TTTTTTTTTTTTTGAGACAGAG-TCTCGCTCTTGTCTCCCAGGCTGGAGGGCAATGCGCGA 3150

Qy   61 tctcagctcaccgcaacctccgcctcccgggttcaagcgatt--ctcctcgcctcagcctc 118
             ||| |||||||| || |||||||||||||||||||||||| | | ||||||||||||||
Db   3150 ...   3150

```
Db  3149 TCTTCGGCTCACTGCAACCTCCCGGGTTCAGTGATTTCCCATGCCTC 3090

Qy   119 cccagtagctgggattacaggcgcatgcaccaccgctcggctaatttgtattttttt 178
         | ||||||||||||||| |  || ||| || ||||| |||||  ||| |||||| ||
Db  3089 CCGAGTAGCTGGGATTATAGTTGCGTGCCACCACCTGGCTAACTTT---TGTATTTTT 3033

Qy   179 agtagagatggagttctcatgttggtcaggctggtctcgaactccgagacctcagatga 238
         ||||| | || |||||||| |||||||||||||| | |||||||| || |||||| ||
Db  3032 AGTAGAAACGGGGTTTCACCATGTTGGCCAGGCT-GTCTTGAACTCCTGACCTCAGGTGA 2974

Qy   239 tccctcccgtctcggcctcccaaagtgctag---atacaggactggccaccatgcccggc 294
         |||  |  | |||| |||||||||||||| ||   |||||| ||| ||||| | || |||
Db  2973 TCCACCTGCCTTGGCCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCGCCCAGCT 2914

Qy   295 tctgcctggctaattttttgggtagaaacaggggtttcactgatgtgcccaagctggtctc 354
         ||||  | ||| |||||| | |  ||||||  ||||||| | | ||| ||| |||||| ||
Db  2913 AATAAAAATAATAATAATTATTTTTTTGAGACGGAGTCTCGCTCTGTCGCCCAGGCTGGAGTG 2854

Qy   355 ctgagctcaagtcaactgtcctcagccgctcccaaagtgctgggattacaggcgtgcag 414
         |||       |  | | | | | ||||| |||| |||||| ||| |||||| || |||
Db  2853 CGGA---CTGCAGTGGCGCAAATCTCGGCTCACTGCAAGCTCCGCTTCCCGGGTTCAGC 2798

Qy   415 ccgtgcctggccttttatttttttaagacacaggtgtcccactcttacccagg 474
         |  | ||  ||| || ||  |||| | | ||||| | |||||| || |||| |
Db  2797 CATTCTCCTGCCTCAGCCTCCTGAGTAGCTGGGACTACAGGCGCCCCACCGCCCAGCT 2738

Qy   475 atgaagtgcagtgtgtgatccag-----------ctcactgcagccttca 515
         || ||||||  | |   |   |           ||||| |  |||| |
Db  2737 CTAATTTTTTGTATTTTTTAGTAGAGACGGGGTTTCACCTTGTTAGCCAGGATGGTCTCGA 2678

Qy   516 actcctgagatcaagcaatcctcctgcctcagcctcccaagtgctgggaccaaaagacatg 575
         |||| || | ||| ||| ||||| |||||| ||||||||||||||||| | | ||| |||
Db  2677 TCTCCCTGACCTCATG-ATCCACCCGCCTCCCAAAGTGCTGGGATTACAGGCGTG 2619

Qy   576 caccactacaactgctaatttttattttttgagacagaagtctcaac 635
         ||||  |   | |||||| |||  ||||||||| | ||| |||| ||
Db  2618 AGCCACCGCGCCCAGC------CTTTTTTTTTTTTTTTTTTGAGACACAGTCTC-AC 2568

Qy   636 tctgtcaccaggctggagtgcagtggcgcaatcttggctcactgcaacctctgcctccc 695
         |||||||| |||||||| |||||||| | |||| | ||||||| |  |||| |||||||
Db  2567 TCTGTCACCCAGGCTGGAATGCAGTGGCCC-ATCTCAGCTCACTGACCTCCACCTCCC 2509

Qy   696 gggttcaagttattcctcctgcctccagcctcgagtagctgggactacaggcgccacca 755
         ||||||||| ||||||||||||| || ||| ||||||||||| |||||||||||||||
Db  2508 GGGGTTCAAGTGATTCTCCTGCCTCAGTCTCCCGAGTAGCTGGGATTACAGGCGCCACCA 2449

Qy   756 cgcctagctaat-ttttgtatttttagagatgggg--ttcaccatgttcgccaggt 813
         | | |  || |  |||||| |||||||| | |||   ||||||| | | |||||
Db  2448 CCATGCCCGGCTCATGTTTGTATTTTTAGGTGGCAGGGGTTTCACCAGGCTGCAGGC 2389

Qy   814 tgatcttgatctctggact--tgtgatctgcctgcctcggcctccaaagtgctgggat 871
         || |||||| |  ||| |   ||||| | |||  || ||||| |||||||||||||||
Db  2388 TGGTCTTGAACTCCTGACCTCAGGTGATCCACCCGCCTTGGCCTTCCAAAGTGCTGGGAT 2329

Qy   872 tacaaggcgtgagccacacgcgcccggct--tattctttaattttgtgtttgaaatggaa 929
         |||| ||||||||||||| | ||||| ||  || | |||| ||||| |||||| ||| |
Db  2328 TACAGGCGTGAGCCACCACCGCGCCCAGCTAATAAAAATAATTATTTTTTTGAGACGGGAG 2269

Qy   930 tctcacttcgttaccacggctggagtgcaatggcgtcactgcaacctctg 989
         ||||  | ||| |||| |||||||||||| ||||| | |  |   | |||
Db  2268 TCTCGCTCTGTCACCAGGCTGGAGTGCAGTGGCGCAAGTCCA 2209

Qy   990 cctccccgggctcaaagcagattctcctgttcagcctcccaaagtggggattacgggcac 1049
         ||||| |||| ||| | || ||||||||| |||||||||| ||||   |||||| | |||
Db  2208 CCTCCCGGGTTCACACATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGCAC 2149

Qy  1050 ctgccacacacccgctaa-ttttgtattcttcattagagggcggggtttcaccatatt 1108
         || ||||| | ||||  | ||||| |||  |  ||||| | |||||||||||| | ||
Db  2148 CCACCACCACGCCCGGCTAATTTTTTGTATTTTTAGTAGAGACGGGGTTTCACTGTGTTA 2089

Qy  1109 gtcaggctggtctccaaactcctgacctcaggtgacccaccctgcctcagcctccaaagtg 1168
         ||||||  |||||| ||| ||||||||||| ||| | ||| || |  |||||||| |||||
Db  2088 GCCAGGGTAGTCTCGATCTCCTGACCTCA-TGATCCGCCCCACTCGGCCTCCCAAAGTG 2031
```

```
Qy  1169 ctgggattacaggcgtgagcccacctcaccagccgcggctaatttagataaaaaatat 1225
         ||||||||||||||||||||||  ||||  || || ||||||  ||| | |  ||||
Db  2030 CTGGGATTACAGGCGTGAGCCACGGCCACCGCGCCCAGCCCATAAAATAATTTTTTAAAAAAT 1974
```

RESULT 10
AAI63343/c
ID   AAI63343/c
XX
DE   AAI63343 standard; DNA; 31853 BP.
XX
AC   AAI63343;
XX
DT   22-OCT-2001 (first entry)
XX
DE   Human kidney related polynucleotide SEQ ID NO 658.
XX
KW   Human; kidney antigen; immunosuppressive; antiarthritic; antirheumatic;
KW   antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;
KW   nootropic; neuroprotective; antibacterial; virucide; fungicide;
KW   opthaimological; antiallergic; hepatotropic; antidiabetic;
KW   antiinflammatory; antiulcer; vulnerary; anticonvulsant; antiparasitic;
KW   gene therapy; cancer; immune disorder; cardiovascular disorder;
KW   neurological disease; infection; ds.
XX
OS   Homo sapiens.
XX
XX
PN   WO200155323-A2.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01343.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
```

PR   05-SEP-2000;   2000US-0229513.
PR   05-SEP-2000;   2000US-0230437.
PR   06-SEP-2000;   2000US-0230438.
PR   06-SEP-2000;   2000US-0231242.
PR   08-SEP-2000;   2000US-0231243.
PR   08-SEP-2000;   2000US-0231244.
PR   08-SEP-2000;   2000US-0231413.
PR   08-SEP-2000;   2000US-0231414.
PR   08-SEP-2000;   2000US-0232080.
PR   08-SEP-2000;   2000US-0232081.
PR   12-SEP-2000;   2000US-0231968.
PR   14-SEP-2000;   2000US-0232397.
PR   14-SEP-2000;   2000US-0232398.
PR   14-SEP-2000;   2000US-0232399.
PR   14-SEP-2000;   2000US-0232400.
PR   14-SEP-2000;   2000US-0232401.
PR   14-SEP-2000;   2000US-0233063.
PR   14-SEP-2000;   2000US-0233064.
PR   14-SEP-2000;   2000US-0233065.
PR   21-SEP-2000;   2000US-0234223.
PR   21-SEP-2000;   2000US-0234274.
PR   25-SEP-2000;   2000US-0234997.
PR   25-SEP-2000;   2000US-0234998.
PR   26-SEP-2000;   2000US-0235484.
PR   27-SEP-2000;   2000US-0235834.
PR   27-SEP-2000;   2000US-0235836.
PR   29-SEP-2000;   2000US-0236327.
PR   29-SEP-2000;   2000US-0236367.
PR   29-SEP-2000;   2000US-0236368.
PR   29-SEP-2000;   2000US-0236369.
PR   29-SEP-2000;   2000US-0236370.
PR   02-OCT-2000;   2000US-0236802.
PR   02-OCT-2000;   2000US-0237037.
PR   02-OCT-2000;   2000US-0237038.
PR   02-OCT-2000;   2000US-0237039.
PR   02-OCT-2000;   2000US-0237040.
PR   13-OCT-2000;   2000US-0239935.
PR   13-OCT-2000;   2000US-0239937.
PR   20-OCT-2000;   2000US-0240960.
PR   20-OCT-2000;   2000US-0241221.
PR   20-OCT-2000;   2000US-0241785.
PR   20-OCT-2000;   2000US-0241786.
PR   20-OCT-2000;   2000US-0241787.
PR   20-OCT-2000;   2000US-0241808.
PR   20-OCT-2000;   2000US-0241809.
PR   20-OCT-2000;   2000US-0241826.
PR   01-NOV-2000;   2000US-0244617.
PR   08-NOV-2000;   2000US-0244474.
PR   08-NOV-2000;   2000US-0244475.
PR   08-NOV-2000;   2000US-0244476.
PR   08-NOV-2000;   2000US-0244477.
PR   08-NOV-2000;   2000US-0244478.
PR   08-NOV-2000;   2000US-0245523.
PR   08-NOV-2000;   2000US-0245524.
PR   08-NOV-2000;   2000US-0245525.
PR   08-NOV-2000;   2000US-0245526.
PR   08-NOV-2000;   2000US-0245527.
PR   08-NOV-2000;   2000US-0245528.
PR   08-NOV-2000;   2000US-0245532.
PR   08-NOV-2000;   2000US-0246609.
PR   08-NOV-2000;   2000US-0246610.
PR   08-NOV-2000;   2000US-0246611.
PR   08-NOV-2000;   2000US-0246613.
PR   17-NOV-2000;   2000US-0249207.
PR   17-NOV-2000;   2000US-0249208.
PR   17-NOV-2000;   2000US-0249209.
PR   17-NOV-2000;   2000US-0249210.
PR   17-NOV-2000;   2000US-0249211.
PR   17-NOV-2000;   2000US-0249212.
PR   17-NOV-2000;   2000US-0249213.
PR   17-NOV-2000;   2000US-0249214.
PR   17-NOV-2000;   2000US-0249215.
PR   17-NOV-2000;   2000US-0249216.

PR   17-NOV-2000;   2000US-0249217.
PR   17-NOV-2000;   2000US-0249218.
PR   17-NOV-2000;   2000US-0249244.
PR   17-NOV-2000;   2000US-0249245.
PR   17-NOV-2000;   2000US-0249264.
PR   17-NOV-2000;   2000US-0249265.
PR   17-NOV-2000;   2000US-0249297.
PR   17-NOV-2000;   2000US-0249299.
PR   17-NOV-2000;   2000US-0249300.
PR   01-DEC-2000;   2000US-0250160.
PR   01-DEC-2000;   2000US-0250391.
PR   05-DEC-2000;   2000US-0251030.
PR   05-DEC-2000;   2000US-0251988.
PR   06-DEC-2000;   2000US-0256719.
PR   08-DEC-2000;   2000US-0251479.
PR   08-DEC-2000;   2000US-0251856.
PR   08-DEC-2000;   2000US-0251868.
PR   08-DEC-2000;   2000US-0251869.
PR   08-DEC-2000;   2000US-0251989.
PR   08-DEC-2000;   2000US-0251990.
PR   11-DEC-2000;   2000US-0254097.
PR   05-JAN-2001;   2001US-0259678.

PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA, Barash SC, Ruben SM;
XX

DR   WPI; 2001-488784/53.
XX

PT   New isolated nucleic acids and polypeptides, useful for diagnosing,
PT   treating and/or preventing human diseases and disorders -
XX

PS   Disclosure; SEQ ID NO 658; 564pp + Sequence Listing; English.
XX

CC   The invention relates to novel kidney related polynucleotides
CC   (AAI62971-AAI63793) and the encoded polypeptides (AAM42417-AAM42691)
CC   collectively known as kidney antigens and the use of such kidney antigens
CC   for detecting disorders of the kidney, especially kidney cancer and
CC   kidney cancer metastases. The polynucleotides and proteins are also
CC   useful for preventing, treating or ameliorating medical conditions
CC   e.g. by protein or gene therapy. The genes are isolated from a range
CC   of human tissues disclosed in the specification. The nucleic acids,
CC   proteins, antibodies and (ant)agonists are useful in the diagnosis,
CC   treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer,
CC   and other cancers of the adrenal gland, bone, bone marrow, breast,
CC   gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders
CC   e.g. Addison's disease, allergies, autoimmune haemolytic anaemia,
CC   autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple
CC   sclerosis, rheumatoid arthritis and ulcerative colitis;
CC   (c) cardiovascular disorders such as myocardial ischaemias; (d) wound
CC   healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;
CC   and (f) infectious diseases such as viral, bacterial, fungal and
CC   parasitic infections.
CC   Note: The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX

SQ   Sequence 31853 BP; 8092 A; 6947 C; 7314 G; 9500 T; 0 other;

Query Match   33.2%; Score 478.2; DB 22; Length 31853;
Best Local Similarity   70.8%; Pred. No. 5.3e-117;
Matches 890; Conservative   0; Mismatches 313; Indels   54; Gaps   17;

```
Qy          1 tttttttttttgagatggagtttcgctcttgttgcccaaggctggagtgcaatggcgcaa   60
              ||||||||||||| ||      ||| ||| ||||||||||||||||||||||||||| | 
Db       3208 TTTTTTTTTTTTGAGACAGAG-TCTCGCGTCTTGTCTCCCAGGCTGGAGGCAAATGCGCGA 3150

Qy         61 tctcagctcaccgcaacctccgcctcccgggttcaaacggatt--ctcctgcctcagcctc  118
              ||  ||| ||||| |||||| ||||| ||||||||| ||||||  |||||||||||||||| 
Db       3149 TCTTGGCTCACTGCAACCTCCACCTCCCGGGTTCAAGTGATTCTCCTGCCTCAGCCTCCTC 3090

Qy        119 cccagtagctgggattacaggcatgtgccaccacgctcggctaatttttgtattttttt  178
```

```
Db   3089 CCGAGCTAGCTGGGGATTATAGTTGCGTTGCCACCACCTGGCTAACTTT---TGTATTTTT 3033
          | || ||||||||||||||||||||||||||||||||||||||||||||   ||||||||
Qy    179 agtagagatgggagtttctccatgtgttggtgtgctggtctcgaacctccgacctcagatga 238

Db   3032 AGTAGAAACGGGGTTTCACCANGTTGGCCAGGCT-GTCTTGAACTCCTGACCTCAGGTGA 2974
          ||||| | |  |||||||||| ||| |||||||| |  | ||||||||||||||||||||
Qy    239 tccctccgtcctcggcctcccaaagtgctag---atacaggactggcaccatgccggc 294

Db   2973 TCCACCTGCCTTGGCCTCCCAAAGTGCTGGGATTACAGGCGTGCACCACCGCGCCCAGCT 2914
          ||| |  ||||||||||||||||||||||||||||||||| |||||||| |||||||||
Qy    295 tctgcctggctaatttttgtggtagaaacaggggtttcactgatgtgcccaagctggtctc 354

Db   2913 AATAAAATAATAATTATTTTTTTTTTGAGACGGAGTCTCGCTCTGTCGCCCAGGCCGGACTG 2854
          || | |||| | ||| ||||||||||||||||||||||||||||||||||||||||||||
Qy    355 ctgagctcaagcagtccacctgcctccagcctcccaaagtgctgggattacaggcgtgcag 414

Db   2853 CGGA----CTGGCAGTGGCGGCAAATCTCGGCCTCACTGCAAGCTCCGCTTCCCGGGTTCACGC 2798
          | ||    ||||| ||||||||||||||||||||||||||||||||||||||||||||||
Qy    415 ccgtgcctggcctttattttttaagacacaggtgtcccaacttcttaccaagg 474

Db   2797 CATTCTTCCTGCCTCTGACTGAGTCCAGGCTGGGACTACAGGCGCCCACCGCGCCCGG 2738
          ||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    475 atgaagtgcagtgggtgtgtgatcacag--------------cttcactgcagccttca 515

Db   2737 CTAATTTTTTGTATTTTTAGTAGAGACGGGGTTTCACCTTGTTAGCCAGGAGGTCTCGA 2678
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    516 actcctgagga tcaagcaatcctctgcctcagcctcccagagtagctgggattacagg 575

Db   2677 TCTCCTGACCTCATG-ANCCACCGCGCCTCCCAAAGTGCTGGGATTACAGGCGTG 2619
          |||||||||||| |  | |||||||||||||||||||||||||||||||||||||
Qy    576 caccactacactggctaatttttattttttttattttttgagacagagtctcaac 635

Db   2618 AGGCACCCGGCGCCCAGC-------CTTTTTTTTTTTTTTTTTTTTTTGAGACACAGTCTC-AC 2568
          ||||||||||||||||||       ||||||||||||||||||||||||||||||| ||
Qy    636 tctgtcacccaggctgga gtgcagtggcgcaatctggttcttcactgcaacctctgcctcc 695

Db   2567 TCTGTCACCCAGGCTGGAAGTGCAGTGGCCC-ATCTCAGTCTCACTGGAAGTAACAGGCCACCTCCC 2509
          |||||||||||||||||||||||||||||||  |||||||||||||||||||| |||||||
Qy    696 gggttcaagttcctccctgagtagctgggact acaggcgcccacca 755

Db   2508 GGGGTTCAAGTGATTCTCCTGCCTCCAGTCTCCCGAGTAGCTGGGATTACAGGCGCACCCACCA 2449
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    756 cgcctagctaat-ttttgtatttattagga gatgggg-ttcaccatgttcgccaagt 813

Db   2448 CCATGGCCCGGCTCANGTTTGTATTTTTAGTGGAGACGGGGTTTCACCATGTTGGCCAGGC 2389
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    814 tgtatcttgatctctgyaacct--tgtgatctgcctcggcctcccaaagtgctgggat 871

Db   2388 TGGTCTTGAACTCCTGACCTCAGGTGATCCACCTGCCTCAGCCTCCCAAAGTGCTGGGAT 2329
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    872 tacaggcgtgagccaccacgcgccggct--tattttattttttgttttgaaatggaa 929

Db   2328 TACAGGGGTGGGAGCCACCACGCGCCCAGCTAATAATAATAATTATTATTTTTTTTTGAGACGGAG 2269
          |||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Qy    930 tcacctctgttaccaggctggagtgcaaatctcggctcactgcaacctctg 989

Db   2268 TCTCGCTCTGTCACCCAGGCTGGAGTGGTCGCAGTGGCTCACGCTGCAAGGTCCA 2209
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    990 cctccccgggcctcaaagcgattcctcctgcctcagcctgggattacgggcac 1049

Db   2208 CCTCCCCGGTTCACACANTTCTCCTGCCTCCAGCCTCCCAGTAGCTGGGAATTACAGGCAC 2149
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1050 ctgccaccacacccgctaa-tttttgtatttccattagaggggcgggggtttcaacatattt 1108

Db   2148 CCACCCACCACACCCGGCTAATTTTTTGTATTTTTAGTAGAGACGGGGTTTCACCTGTGTTA 2089
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1109 gtcaggctggtctcaaactcctgaccctgacccagcctcccaaagtg 1168

Db   2088 GCCAGGGTAGTCTCGATCTCCTGACCTCA-TGATCCGCCACCTTCGGCCTCCCAAAGTG 2031
          |||||||||||||||||||| |||||||| ||||||||||||||||||||||||||||
Qy   1169 ctgggattacaggcgtgagccacctccagccggctaatttagataaaaatat 1225
```

```
Db   2030 CTGGGATTACAGGCGTGAGCCACCGCGCCCAGCGCCCATAAAATAATTTTTTAAAAAAT 1974
```

Db   2030

RESULT   11

AAS39682/c

ID   AAS39682 standard; DNA, 11319 BP.
XX
AC   AAS39682;
XX
DT   17-DEC-2001 (first entry)
XX
DE   Genomic sequence #101 encoding human colon associated polypeptide.
XX
KW   Human; colon cancer; congenital abnormality; infection; colitis;
KW   inflammatory bowel disease; IBD; neoplastic disorder; gene therapy;
KW   intestinal inflammatory disorder; malabsorption syndrome; gastric;
KW   sigmoid disease; antibacterial; antiviral; antiinflammatory;
KW   cytostatic; ds.
XX
OS   Homo sapiens.
XX
PN   WO200155302-A2.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01240.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0217487.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225566.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   18-AUG-2000; 2000US-0225759.
PR   22-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227109.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229944.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.

```
PR   08-SEP-2000;  2000US-0231244.
PR   08-SEP-2000;  2000US-0231413.
PR   08-SEP-2000;  2000US-0231414.
PR   08-SEP-2000;  2000US-0232080.
PR   08-SEP-2000;  2000US-0232081.
PR   12-SEP-2000;  2000US-0231968.
PR   14-SEP-2000;  2000US-0232397.
PR   14-SEP-2000;  2000US-0232398.
PR   14-SEP-2000;  2000US-0232399.
PR   14-SEP-2000;  2000US-0233400.
PR   14-SEP-2000;  2000US-0232401.
PR   14-SEP-2000;  2000US-0233063.
PR   14-SEP-2000;  2000US-0233064.
PR   14-SEP-2000;  2000US-0233065.
PR   21-SEP-2000;  2000US-0234223.
PR   21-SEP-2000;  2000US-0234274.
PR   25-SEP-2000;  2000US-0234997.
PR   25-SEP-2000;  2000US-0234998.
PR   26-SEP-2000;  2000US-0235484.
PR   27-SEP-2000;  2000US-0235834.
PR   27-SEP-2000;  2000US-0235836.
PR   29-SEP-2000;  2000US-0236327.
PR   29-SEP-2000;  2000US-0236367.
PR   29-SEP-2000;  2000US-0236368.
PR   29-SEP-2000;  2000US-0236369.
PR   29-SEP-2000;  2000US-0236370.
PR   13-OCT-2000;  2000US-0239937.
PR   29-SEP-2000;  2000US-0240960.
PR   20-OCT-2000;  2000US-0241221.
PR   20-OCT-2000;  2000US-0241785.
PR   20-OCT-2000;  2000US-0241786.
PR   20-OCT-2000;  2000US-0241787.
PR   20-OCT-2000;  2000US-0241808.
PR   20-OCT-2000;  2000US-0241809.
PR   20-OCT-2000;  2000US-0241826.
PR   01-OCT-2000;  2000US-0244617.
PR   08-NOV-2000;  2000US-0244474.
PR   08-NOV-2000;  2000US-0244475.
PR   08-NOV-2000;  2000US-0244476.
PR   08-NOV-2000;  2000US-0244477.
PR   08-NOV-2000;  2000US-0244478.
PR   08-NOV-2000;  2000US-0246523.
PR   08-NOV-2000;  2000US-0246524.
PR   08-NOV-2000;  2000US-0246525.
PR   08-NOV-2000;  2000US-0246526.
PR   08-NOV-2000;  2000US-0246527.
PR   08-NOV-2000;  2000US-0246528.
PR   08-NOV-2000;  2000US-0246532.
PR   08-NOV-2000;  2000US-0246609.
PR   08-NOV-2000;  2000US-0246610.
PR   08-NOV-2000;  2000US-0246611.
PR   08-NOV-2000;  2000US-0246613.
PR   17-NOV-2000;  2000US-0249265.
PR   17-NOV-2000;  2000US-0249297.
PR   17-NOV-2000;  2000US-0249299.
PR   17-NOV-2000;  2000US-0249300.
PR   01-DEC-2000;  2000US-0250160.
PR   01-DEC-2000;  2000US-0250391.
PR   05-DEC-2000;  2000US-0251030.
PR   05-DEC-2000;  2000US-0251988.
PR   05-DEC-2000;  2000US-0256719.
PR   06-DEC-2000;  2000US-0251479.
PR   08-DEC-2000;  2000US-0251855.
PR   08-DEC-2000;  2000US-0251856.
PR   08-DEC-2000;  2000US-0251868.
PR   08-DEC-2000;  2000US-0251869.
PR   08-DEC-2000;  2000US-0251989.
PR   08-DEC-2000;  2000US-0251990.
PR   11-DEC-2000;  2000US-0254097.
PR   05-JAN-2001;  2001US-0259678.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA, Barash SC, Ruben SM;
XX
WPI; 2001-465567/50.
XX
DR
XX
PT   Isolated polypeptide for treating, preventing and/ or prognosing
PT   disorders related to the colon including colon cancers and also for
PT   testing and detection e.g. diagnosis -
XX
PS   Disclosure; SEQ ID NO 579; 562pp; English.
XX
CC   The present invention relates to the isolation of novel human colon
CC   associated polypeptides (AAU22468-AAU22701), and the cDNA and genomic
CC   sequences encoding for them. The sequences of the invention are useful
CC   in the diagnosis, treatment, prevention and/or prognosis of disorders
CC   of the colon including colon cancer, congenital abnormalities
CC   (e.g. atresia and stenosis), bacterial and viral infections,
CC   inflammatory bowel disease (IBD), neoplastic cell disorders
CC   (e.g. polyps and adenomas, intestinal inflammatory disorders, colitis,
CC   colonic inflammation, diarrhoea and dysentery, malabsorption syndromes
CC   (e.g. lactose intolerance), intestinal obstruction and sigmoid diseases.
CC   The polynucleotides sequences of the invention can also be used in gene
CC   therapy. AAS39582-AAS40060 represent DNA sequences encoding for the
CC   novel human colon associated polypeptides of the invention.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences..
XX
SQ   Sequence 11319 BP; 3454 A; 2363 C; 2754 G; 2748 T; 0 other;

Query Match             33.1%;  Score 477.4;  DB 22;  Length 11319;
Best Local Similarity   74.8%;
Matches  699;  Conservative    0;  Mismatches  211;  Indels   25;  Gaps    7;

Qy         350 gtctcctgagctcaagcagtcctacctgcctcagcctcccaaagtgctggattacaggcg 409
               |||||| |||| |     | ||||||||||||||||||| |||||||||||||| ||||
Db        1222 GCCTCCCGTGTTCAAGTGATTCTCGTGCCTCAGCCTTCGAGTAGCTGGGATTACGGGGCG 1163

Qy         410 tgca---------gccgtgcctggcctttattttattttttaagacacaggtgtgtcc 461
               ||||         || | ||| ||  || ||  ||  |||||||||||| |  ||||
Db        1162 TGCACCACCACGGCGTGGTTAATTTCTTTCTTTTTTTTTTTTTTAAGGCCAAGTCTCGC 1103

Qy         462 actcttacccaggatgaagtgcagtggtgtgatcacagctcactcagccttcaactct 521
               |||   |  |||||| ||  || ||| ||||   | ||  || |||||||  |||  ||
Db        1102 TCTGTGTCACCCAGGCTGCAGTGCCAGTGGTGCCACTCACTTGCGCACCTCTGCCTCCA 1043

Qy         522 gagatcaagc-atcctcctgcctcagcctcccaagtagctgggaccaaaagacatgcacca 580
               | ||||||||  ||||  ||||||||||||| | ||||||||||  |  ||||||||||| |
Db        1042 GGGTTCAAGCGAATTCTCCCACCTCAGCCTCTCGGGTAGCTGGGAATTACAGGCACATGCCA 983

Qy         581 ctacacctggctaattttttatttttaat-----------ttttttgagacagagtct 631
               ||| | | ||||||||||| |  |||| |           ||    ||||||| |||||
Db         982 CTACGCCCAGGCTAATTTTTGTTTTCTGTTTTTTTTTTGTTTGCTTGTTTTTTGAGACGGAGTCT 923
```

```
RESULT 12
AAK90025/c
ID    AAK90025 standard; DNA; 11319 BP.
XX
AC    AAK90025;
XX
DT    05-NOV-2001  (first entry)
XX
DE    Human digestive system antigen genomic sequence SEQ ID NO: 3601.
XX
KW    Human; digestive system antigen; gene therapy; cancer; appendicitis;
KW    ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
KW    digestive system disorder; Meckel's diverticulum; ds.
XX
OS    Homo sapiens.
XX
PN    WO200155314-A2.
XX
PD    02-AUG-2001.
XX
PF    17-JAN-2001; 2001WO-US01324.
XX
PR    31-JAN-2000; 2000US-0179065.
PR    04-FEB-2000; 2000US-0180628.
PR    24-FEB-2000; 2000US-0184066.
PR    02-MAR-2000; 2000US-0186350.
PR    16-MAR-2000; 2000US-0189874.
```

```
Qy    632 caacttctgtcacccaggctggagtgcagtggcgcaatctcggctcactgcaacctctgcc 691
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db    922 C-GCTGTGCCCCAGGCTGGAGTGCAGTGGCGCAATCTCGGCTCACTGCAAGCTCCGCC 864

Qy    692 tcccggggttcaagttattctcctgcccagcctcctgagtagctgggactacaggcgcc 751
          ||| || ||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db    863 TCCCGGGTTCACGCATTCTCCTGCCTCAGCCTCCCGATTGGCTAGGACTACAGGCGCCC 804

Qy    752 accac--gcctagctaatttttgtatttttagtagagatgggg--ttcaccatgttcgc 808
          |||||  || ||  |||||| ||||||||||||||||||||| ||||||||||||||||
Db    803 GCCACCATGCCCCGGCTAATTTTTGTATTTTTAGTAGAGACGGGGTTTTACCGTGTTAGC 744

Qy    809 caggttgatctctggacccttgtgatctgcctcggcctcccaaagtgctgg 868
          ||| || ||||||||||||||||||||||||||||||||||||| |||| |
Db    743 GAGGATGGTCTCGATCTCCTGACCTTGTGATCTGCCCGCCTCCCAAAAGTGCTGG 684

Qy    869 gattacaggcgtgagccacacgcccggcttattctgtttttattttttgttt---gtttgaaat 925
          |||||| ||| ||||||||||||||||||||||||| | |||||||||||| ||||||||
Db    683 GATTACAGGCGTGAGCCACCACTGCACCCGGCCTTGGTTTTGTTTGTTTGTTTGGTTTTGAGAC 624

Qy    926 ggaatctcactgttacccaggctgggagtgcaatggcgcaaatctcggctcactgcaacc 985
          |||| ||||||||||||||||||||||||||||||||||||||||||| ||||||
Db    623 GGAGTCTCACTCTGTCACCCAGGCTGGAGTGCAGTGGCGCGGATCTCAGCTCTCTGCAACC 564

Qy    986 tctgcctcccgggtcaagcgaattctcctgtctccagcctcccaagcagctgggattacgg 1045
          |||||||| |||||||||| || | |||||||||||||||||||||||||||||||| |||
Db    563 TGTGCCTCCCAGGGTTCAAGTGATTCTCCTGTCTCAGCCTCCCAGTAGCTAGGACCACAG 504

Qy    1046 gcacctgccacccacccccgctaattttttgtattttcattagagacggggtttcaccata 1105
          |||||||| ||||||||||||||||||||||| |||| ||||||||||||||||||||||
Db    503 GTGCACACCACCACGCAAGCTAAATTTTGTATTTATTAGTAGAGATGGGGTTTCACCATG 444

Qy    1106 tttgtccagctggtctcaaactcctgacctcaggtgacccacctgcctcagcccttccaaa 1165
          ||||||| ||||||||||||||||||||| |||||| ||||||||| ||||||||||||
Db    443 TTGGCAGGCTGGTCTCGAAACTCCTGACCTCAAGTGATCTGCCCGCCTCCCGCCTCCCAAA 384

Qy    1166 gtgctgggattacaggcgtgagcgcaccacgcccggcctaatttagataaaaaaatat 1225
          || || ||||||||||||||| |||||||||||||||||| |||||||||||||||
Db    383 GTGCTGGGATTACAGGCGTGAGCCACCATGCCCGGCCAAGTTCTAGCATCTTAAGAAAAC 324

Qy    1226 gtagcaatggggggtcttgctatgtgtgcccaggct 1260
          ||| ||||||||||||||||||||| |||||||||
Db    323 AACTCATGGAAGGGGAGTGAATCTGCACTCCCTGCT 289
```

```
PR   13-OCT-2000;  2000US-0239937.
PR   20-OCT-2000;  2000US-0240960.
PR   20-OCT-2000;  2000US-0242021.
PR   20-OCT-2000;  2000US-0241785.
PR   20-OCT-2000;  2000US-0241786.
PR   20-OCT-2000;  2000US-0241787.
PR   20-OCT-2000;  2000US-0241808.
PR   20-OCT-2000;  2000US-0241809.
PR   20-OCT-2000;  2000US-0241826.
PR   01-NOV-2000;  2000US-0244617.
PR   01-NOV-2000;  2000US-0246474.
PR   08-NOV-2000;  2000US-0246475.
PR   08-NOV-2000;  2000US-0246476.
PR   08-NOV-2000;  2000US-0246477.
PR   08-NOV-2000;  2000US-0246478.
PR   08-NOV-2000;  2000US-0246523.
PR   08-NOV-2000;  2000US-0246524.
PR   08-NOV-2000;  2000US-0246525.
PR   08-NOV-2000;  2000US-0246526.
PR   08-NOV-2000;  2000US-0246527.
PR   08-NOV-2000;  2000US-0246528.
PR   08-NOV-2000;  2000US-0246532.
PR   08-NOV-2000;  2000US-0246609.
PR   08-NOV-2000;  2000US-0246610.
PR   08-NOV-2000;  2000US-0246611.
PR   08-NOV-2000;  2000US-0246613.
PR   17-NOV-2000;  2000US-0249207.
PR   17-NOV-2000;  2000US-0249208.
PR   17-NOV-2000;  2000US-0249209.
PR   17-NOV-2000;  2000US-0249210.
PR   17-NOV-2000;  2000US-0249211.
PR   17-NOV-2000;  2000US-0249212.
PR   17-NOV-2000;  2000US-0249213.
PR   17-NOV-2000;  2000US-0249214.
PR   17-NOV-2000;  2000US-0249215.
PR   17-NOV-2000;  2000US-0249216.
PR   17-NOV-2000;  2000US-0249217.
PR   17-NOV-2000;  2000US-0249218.
PR   17-NOV-2000;  2000US-0249244.
PR   17-NOV-2000;  2000US-0249245.
PR   17-NOV-2000;  2000US-0249264.
PR   17-NOV-2000;  2000US-0249265.
PR   17-NOV-2000;  2000US-0249297.
PR   17-NOV-2000;  2000US-0249299.
PR   17-NOV-2000;  2000US-0249300.
PR   01-DEC-2000;  2000US-0250160.
PR   01-DEC-2000;  2000US-0250391.
PR   05-DEC-2000;  2000US-0251030.
PR   05-DEC-2000;  2000US-0251988.
PR   05-DEC-2000;  2000US-0256719.
PR   06-DEC-2000;  2000US-0251479.
PR   06-DEC-2000;  2000US-0251856.
PR   08-DEC-2000;  2000US-0251868.
PR   08-DEC-2000;  2000US-0251869.
PR   08-DEC-2000;  2000US-0251989.
PR   08-DEC-2000;  2000US-0251990.
PR   11-DEC-2000;  2000US-0254097.
PR   05-JAN-2001;  2001US-0259678.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA, Barash SC, Ruben SM;
XX
DR   WPI: 2001-502630/55.
XX
PT   Polynucleotides encoding digestive system antigens, useful for
     diagnosing, treating, preventing and/or prognosing disorders of the
     digestive system, particularly cancer and cancer metastases -
XX
PS   Disclosure: SEQ ID NO 3601; 986pp; English.
XX
CC   The present invention provides the protein and coding sequences of a
     number of human digestive system antigens. These can be used in the
```

```
CC   diagnosis, treatment and prevention of digestive system disorders,
CC   including cancer, Meckel's diverticulum, bacterial or parasitic
CC   infections, appendicitis, Hirschsprung's disease, chronic colitis or
CC   ulcerative colitis. The present sequence is a genomic DNA fragment
CC   encoding a digestive system antigen of the invention.
XX
SQ   Sequence 11319 BP; 3454 A; 2363 C; 2754 G; 2748 T; 0 other;

Query Match             33.1%;  Score 477.4;  DB 22;  Length 11319;
Best Local Similarity   74.8%;  Pred. No. 5.7e-117;
Matches  699;  Conservative    0;  Mismatches  211;  Indels  25;  Gaps  7;

Qy    350 gtctcctgagctcaagcagtccacctgcctccagcctcccaaagtgctgggattacaggcg 409
          ||||| ||| ||  |||||||||| ||||||||   |||  ||||  |||||||||| ||||
Db   1222 GCCTCCGGTGTTCAAGTGATTCTCGTGCCTCAGCCTTCAGTAGCTGGGATTACGGGGCG 1163

Qy    410 tgca------gccgtcggccctttt-tattttatttttttttaagacacaggtgtcc 461
          ||||      ||| |||||   |||  |||| |||  ||| || |||  |  |||  |||
Db   1162 TGCACCACCAGCGCTGGTTAATTTCTTTCTTTATTCTTTTTTTTTTTTAAGGCCAAGTCTGC 1103

Qy    462 actcttacccaggatgaagtgcagtggtgtgatcaacagctcactgcaagccttcaactct 521
          | ||| ||||||   |||  ||||||||  |||||||| |||||| || |||||| | ||||
Db   1102 TCTGTCACCAGGCTGGAGTGCAGTGGTGCCGACCTTGGCTCACTGCAACCTCTGCCTCCA 1043

Qy    522 gagatcaagc-atcctcctgcctcagcctcccaagtagctgggaccaaagactgcacca 580
          |||||||||  |||||||||||||||| |||||||||||||| |  || ||  |||||||
Db   1042 GGGTTCAAGCGATTCTCCCACCTCAGCCTCCCGGGTAGCTGGAATTACAGGCACATGCCA 983

Qy    581 ctacacctggctaattttatttatttttttttttaat------ttttgagacagagtct 631
          |||| |||||||| |||| |||  ||||  ||| |||       ||||||||| ||||||
Db    982 CTAGGCCAGGCTAATTTTTGTTTTTTTCGTGTTTTTTGCTTGTTGAGACGGAGTCT 923

Qy    632 caactcctgtcaccaggctggagtgcaatcttggctcactgcaacctctgcc 691
          |||| ||||||||||||||||||||||| |||||||||||| ||||| || |
Db    922 C-GCTCTGTCGCCCAGGCTGGAGTGCAGTGGCGCTTCACTGCAAGCTCCGCC 864

Qy    692 tccggttcaagttattctcctgcccccagcctcctgagtagctgggaccacaggcgccc 751
          |||||||||||  |||||||||  |||||||||||||||||||| ||  ||||  |||||
Db    863 TCCCGGGTTCACGGCCATTCTCCTGCCTCAGCCTCCTGAGTAGCTGGGACTACAGGCGCCC 804

Qy    752 accac--gcctagctaattttttttgtattttttagtagagatgggg-ttcaccatgttcgc 808
          ||||    ||| ||| ||||||| || ||||||||||||||||| ||   |||||  ||| ||
Db    803 GCCACCATGCCCGGCTAATTTTTGTATTTTTTAGTAGAGACGGGGTTTACCGTGTTAGC 744

Qy    809 caggttgatcttgacctttgatctgctggcctcccaaagtgctgg 868
          ||||||||| |||||||| |||||| |||||||||||| |||| |
Db    743 GAGGATGGTCTCGATCTCCTGACCTTGTGATCTGCCCACCTCCAAAGTGCTGG 684

Qy    869 gattacaggcgtgagccaccacgcccggcttattttattttgttt---gtttgaaat 925
          ||||||||| |||| ||||||| ||||||||||| || ||| ||||   |||||||||
Db    683 GATTACAGGCATGAGCCACTCACCCGGCCTTGGTTTGTTTTGTTTTTGTTTTGAGAC 624

Qy    926 ggaatctcactgttacccaggctggagtgcaatggcaaatcttggctcactgcaacc 985
          |||||| |||||||| ||||||| |||||| ||| || |||| |||||||| |||||
Db    623 GGAGTCTCACTCTGTCACCCGGCTGGAGTGCAGTGGCGCGATCTCAGCTCTCTGCCAACC 564

Qy    986 tctgcctccgggctcaagcgatctctcctgcctcagcctcccaagctgggattacgg 1045
          |||||||||| |||| ||| |||||||||| ||||||||||||||| |||||||||||
Db    563 TCTGCCTTCCAGGGTTCAAGCGGTTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGACCACAG 504

Qy   1046 gcacctgccacccaccccgctaattttttgtattttagtagcggggtttcaccata 1105
          ||  |||||  |||||||| ||||||||||||||| ||||| ||||||||||||||
Db    503 GTGCACACCACCACACCCGCTAATTTTTGTAGTAGAGATGGGGTTTCACCATG 444

Qy   1106 tttgtcaagctggtctcaaactcctgacctcaggtgacccaccttccaa 1165
          |||||||||||| |||| ||||||||| || |||||||  |||||| |||
Db    443 TTGGCCAGGCTGGTCTCGAACCTCCTGACCTCAAGTGATCCGCCCGCCTCAGCCTCCCAAA 384

Qy   1166 gtgctgggattacaggcgtgagccaccacccaccggctaatttagataaaaaatat 1225
          ||||| ||||||||||| || |||||||| || ||| ||||||| |||| |||||
Db    383 GTGCTGGGATTACAGGGGTGAGCCACCACATGCCTGGCCACAGTTCTAGCCATCTTAAGGAAAC 324
```

Qy   1226 gtagcaatggggggtcttgctatgttgccaaggct 1260
           || | ||| ||||| |||
Db   323 AACTCATGGAAGGGAGTAGATCTGCACTCCCTGCT 289

RESULT 13

ID   AAV27941
AC   AAV27941 standard; cDNA; 14796 BP.
XX
AC   AAV27941;
XX
OS   Homo sapiens.
XX
DE   Survivin gene.
XX
DT   25-SEP-1998 (first entry)
XX
FH   Key   Location/Qualifiers
FH
FT   CDS   2811..12044
FT   /*tag= a
FT   /product= Survivin
FT   exon   2811..2921
FT   /*tag= b
FT   /number= 1
FT   intron   2922..3173
FT   /*tag= c
FT   /number= 1
FT   exon   3174..3283
FT   /*tag= d
FT   /number= 2
FT   intron   3284..5157
FT   /*tag= e
FT   /number= 2
FT   exon   5158..5275
FT   /*tag= f
FT   /number= 3
FT   intron   5276..11954
FT   /*tag= g
FT   /number= 3
FT   exon   11955..12044
FT   /*tag= h
FT   /number= 4
XX
PN   WO9822589-A2.
XX
PD   28-MAY-1998.
XX
PF   20-NOV-1997; 97WO-US21880.
XX
PR   20-NOV-1997; 97US-0975080.
PR   20-NOV-1996; 96US-0031435.
XX
PA   (UYYA) UNIV YALE.
XX
PI   Altieri DC;
XX
DR   WPI; 1998-312475/27.
DR   P-PSDB; W614359.
XX
PT   Modulating apoptosis by controlling the Survivin gene - useful for
PT   treating transplant rejection, degenerative disorders and tumours
XX
PS   Disclosure; Fig 10; 108pp; English.
XX
CC   The survivin gene can be used to control apoptosis through modification
CC   of the gene. Survivin peptides can be used to inhibit cellular
CC   apoptosis, e.g. for enhancing the viability of organs and tissues prior
CC   to their transplantation, for preserving the growth of cells in culture
CC   or for treating conditions involving abnormal apoptosis,

CC   survivin; apoptosis; cellular apoptosis; transplantation; ss;
KW   motor neuron degenerative disease; HIV infection; immunosuppression;
KW   gastrointestinal perturbations; cardiovascular disorder.

CC   e.g. degenerative diseases such as motor neuron degenerative diseases,
CC   HIV infection, dermatological effects of ageing, disorders and diseases
CC   such as immunosuppression, gastrointestinal perturbations, cardiovascular
CC   disorders, apoptosis related to reperfusion damage, rejection of tissue
CC   transplantation and Alzheimer's disease. Agents which block Survivin
CC   activity can be used to treat e.g. tumours.
XX
SQ   Sequence 14796 BP; 3482 A; 3488 C; 3888 G; 3938 T; 0 other;

Query Match   33.0%; Score 475.4; DB 19; Length 14796;
Best Local Similarity   69.8%; Pred. No. 2.2e-116;
Matches 877; Conservative   0; Mismatches 321; Indels   59; Gaps 15;

Qy   1 tttttttttttgagatggagtttcgctcttgttgcccaggctggagtgcaatggcgcaa 60
           ||| ||||||| |||||| ||||||| ||||| ||||||||||||||| || |||||||
Db   5439 tcttggctcactgcaacctctgcctcccgggttcaagcgattcctcctgcctcagcctcc 5380

Qy   61 tctcaagctcaccgcaacctccgcctcccgggttcaagcgaattctcctgcctcagcctcc 120
           |||| ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5380 tcttttttcctgaggtagag-tttcactccttgttcgcccaggctggagtgcaagtggtgca 5438

Qy   121 cagtagtggattacaggcatgtgccaccagctcgggtcaatgtcctaatttgtatttttag 180
           ||||||||||||||||||||| ||||||| || ||||| ||||| ||| || |||||||||
Db   5499 aagtaaactgggattacaggcgtgagccaccacgcccggctaatttt--tgtattttttag 5555

Qy   181 tagagatggagtttctccatgttggtcaggctggtctcgaactccccgacctcaagatgatc 240
           ||||||||||||||| |||||| || |||||||||||||||| |||| ||||| ||| |||
Db   5556 tagagatggggtttcaccacattacccaggctggtctggaactcctgacctc--gtgatt 5613

Qy   241 cctccgtcctcggcctcccaaagtgctagatacaaggaggcgccacatgcccggctctgcc 300
           |||||||||||||||||||||||||||| ||| || ||| ||||| ||||||||||||||
Db   5614 cgcccaccttggcctcccaaagtgctgg--gattacaggcgtgaaccaccgcggctggc 5670

Qy   301 tggctaattttttgtgtagaacaggggtttcactgatgtgcccaagctgtctctg-- 357
           | |||||| | |||||||||||||||||||| ||| ||||||||||||||||||||
Db   5671 ttttttttttgttct-gagacacagtttcactttaccccagctggagtgggtgg 5729

Qy   358 --agcttcaaagcagtcacctgcctcagcctcccaaagtgctggattacaggcgtgagc 415
           |||||||| ||| |||||||| ||| |||||||||||| ||||||||||||||| |||
Db   5730 cctgatctcggatcactgcaacctctgggctcccaagtgattgcctgcttcagc 5789

Qy   416 cgtgcctgcccttttttatttttttaagacacaggtgtcccactcttaccaagga 475
           | |||||||||||||||||| |||||| ||||||||| | |||||||||||||
Db   5790 ctcccaagtagccgagattacaggcatgtgccacacccaggtaatttttgtatttt 5849

Qy   476 tgaagtgcagtgtgtgatca--cagctcactgcaggcctcaactcctgagatcaag-ca 532
           ||||||||| ||| | ||||| ||||| |||||||||| ||| |||| |||||||| ||
Db   5850 ggtagagacgggagtttcaccatgttggccaggctggtctcgaactcctcaggtga 5909

Qy   533 tcctcctgcctcagcctcccaagtagctgggactacaggcgc---ccacacg 591
           ||| | ||||||||||| |||||||||||||| ||| ||||| ||||||||
Db   5910 tccaccccgcctcggcctcccaaagtgctgagattatagggtgagcccaccac 5969

Qy   592 -----taatttttatttttttgagacagagtctcaactgt 640
           |||||||||||||| | ||||||||||||||||||||||
Db   5970 tcaggaagtatttttattttaaattttattatttgagatggagtct-tgctctgt 6028

Qy   641 cacccaggctggagtgcagtggcgcaatcttggctcactgcaacctccgggtt 700
           |||||||||||||||| | || || |||||||||||||||||||| ||||||
Db   6029 cgcccaggctagagtgcagtggcgacggggtctcggtcactggagtctccgggtt 6088

Qy   701 caagtgattctcctgcctcagcctcccgagtagctgggactacaggcgc---ccacacg 757
           ||||| ||||||||||||||||||||||||||| ||||| ||| ||||| |||||||
Db   6089 caagccaattctcctgcctcagcctcccgagtagctggaactacaggcgcccacca 6148

Qy   758 cctagctaatttttgtatttttagtagagacgg-gtttcaccatgttcgccaggttga 816
           ||| |||| ||||||||| |||||| |||| ||| |||||||| ||||||||||||||||
Db   6149 cccgctaattttttgtatttttagtagagacggggtttcaccgtgttagccaggaggg 6208

Qy   817 tcttgatctctggacttgtgatctgcctcggcctcccaaagtgctgggattacag 876
           ||||||||| |||||||| ||||||||| || ||||||||||| |||||||||||
Db   6209 tcttgatctcctgacctcgtgatctgcctgcctcggcctcccaaagtgctggggattacag 6268

```
RESULT 14
AAS21523
ID    AAS21523 standard; cDNA; 14796 BP.
XX
AC    AAS21523;
XX
DT    21-NOV-2001  (first entry)
XX
DE    DNA encoding human survivin.
XX
KW    Survivin; human; mouse; cytostatic; antisense oligonucleotide;
KW    hyperproliferative condition; cancer; apoptosis; cytokinesis; ss.
XX
OS    Homo sapiens.
XX
PN    WO200157059-A1.
XX
PD    09-AUG-2001.
XX
PF    30-JAN-2001; 2001WO-US02939.
XX
PR    02-FEB-2000; 2000US-0496694.
XX
PA    (ISIS-) ISIS PHARM INC.
XX
PI    Bennett CF, Ackermann EJ, Swayze EE, Cowsert LM;
XX
DR    WPI; 2001-488863/53.
XX
PT    Novel antisense compounds for modulating the expression of Survivin and
PT    treatment of cancer -
XX
PS    Example 13; Page 74-79; 120pp; English.
XX
```

The invention relates to antisense oligonucleotides targeted to a nucleic
acid molecule encoding human Survivin, where the antisense
oligonucleotide inhibits the expression of human Survivin. These
antisense oligonucleotides are used in the treatment of an animal
suffering from a disease or condition associated with Survivin, e.g. a
hyperproliferative condition such as cancer, and comprises administering
a therapeutically effective amount of the antisense
oligonucleotide so that expression of Survivin is inhibited. The
oligonucleotides can also be used to treat a human suffering from a
disease or condition characterised by a reduction in apoptosis
comprising administering the antisense oligonucleotide to a human. In
addition, the antisense oligonucleotide and a cytotoxic chemotherapeutic
agent e.g. taxol or cisplatin, can be used to modulate apoptosis,

```
Qy    877 ggcgtcaagcacacgcccgcgcttattttgtttgttgaaatggaatctcact 936
          |||||||||||| ||| |||||||||||| ||||||||||
Db   6269 gtgtgagccaccaccccggct------------atttttattttttgagacaggagactact 6320

Qy    937 ctgttaccaggctggagtgcaatggccaaatctcggtctcactgcaactctcgctcccg 996
          ||||||||| ||||||| ||||||| || | | |||||||||||||||||||||| ||
Db   6321 ctgtcacctcgggtcgcgtcagtcg-tacaccatagctcactgcagcctcgaacctcctg 6379

Qy    997 ggctcaagcgagttctcctgtctcagcctcccaagcagctgggattacgggcactgccac 1056
          ||||||| || |||| ||||||||||| |||||||| |||||||||||||| || |||
Db   6380 agctcaagtgatcctcccacctcacaagtaattgggactaacaggtgcacccac 6439

Qy   1057 cacaccccgctaattttgtat----------------------tttcattagaggcgggtttcac 1101
          ||| ||||||||||||||||                     ||||||||| || || |||
Db   6440 catgccaacctaattatttattttattttcatagagatgaggtttccc 6499

Qy   1102 catatttgtcaggtcggtgctcaaaactcctgacctcaggtgacccaactcgctcagcctc 1161
          ||| ||||||||||| ||||||||||||||||||| | ||||| ||||||||||| ||||
Db   6500 tgtgttgttccaggctggtcttgaactcctgagctcaggtgatcctttgcctggggcctcc 6559

Qy   1162 caaagtgctgggattacaggcgtgagcacctcaccacatgatccc taatttagataaa 1218
          ||||||||||||||||||||| |  |||||||| | ||||||| ||||| ||||||||
Db   6560 caaagtgctgagattacaggcgatgagcacacgtgcccagctaggaatcattttaaa 6616
```

```
RESULT 15
AAH47531
ID   AAH47531 standard; DNA; 14796 BP.
XX
AC   AAH47531;
XX
DT   30-NOV-2001 (first entry)
XX
DE   Human Her-3 genomic DNA sequence.
XX
KW   Her-3; epidermal growth factor; EGF; receptor/tyrosine kinase; human;
KW   antiinflammatory; cytostatic; antibacterial; antisense; ds.
XX
OS   Homo sapiens.
XX
XX
FH   Key               Location/Qualifiers
FT   CDS               2811..12044
FT                     /*tag= a
FT                     /note= "contains introns"
FT
XX
PN   US6277640-B1.
XX
PD   21-AUG-2001.
XX
PF   31-JUL-2000; 2000US-0630706.
PR   31-JUL-2000; 2000US-0630706.
XX
PA   (ISIS-) ISIS PHARM INC.
XX
PI   Bennett CF, Cowsert LM;
XX
DR   WPI; 2001-535134/59.
DR   P-PSDB; AAG65360.
XX
PT   Antisense compounds capable of modulating expression of human Her-3, a
PT   member of epidermal growth factor family of receptor/tyrosine kinases,
PT   useful for preventing or delaying infection, inflammation or tumor
PT   formation
PT   -
PS   Example 15; Columns 57-70; 49pp; English.
XX
CC   The invention provides antisense compounds capable of inhibiting the
CC   expression of human Her-3, a member of epidermal growth factor (EGF)
CC   family of receptor/tyrosine kinases. The antisense oligonucleotides are
CC   useful for inhibiting the expression of Her-3 in cells or tissues. They
CC   are commonly used as research reagents and in diagnostics for example, to
CC   elucidate the function of particular genes. The antisense compounds are
CC   also useful for distinguishing between functions of various members of a
CC   biological pathway and for research use. They are also utilized for
CC   diagnostics, therapeutics, prophylaxis and in kits. They are useful
CC   prophylactically, e.g. to prevent or delay infection, inflammation or
CC   tumor formation. The present sequence represents the human Her-3 DNA
CC   sequence (GenBank Accession No. U75285).
XX
SQ   Sequence 14796 BP; 3482 A; 3488 C; 3938 G; 3888 T; 0 other;
```

```
Query Match             33.0%;  Score 475.4;  DB 22;  Length 14796;
Best Local Similarity   69.8%;  Pred. No. 2.2e-116;
Matches  877;  Conservative  0;  Mismatches  321;  Indels  59;  Gaps  15;

Qy      1 tttttttttttgagatggagttctccatgtgttggtcaggctggtctcgaactcccagatc 60
          ||||||||| ||  |||||||||||||| ||||||| ||||||||||||||||||||| |||
Db   5556 tagagatggggtttccaccacattgccaggctggtcttgaactcctgacctc--gtgatt 5613

Qy    241 cctccgttctcggcctcccaaagtgctagatacaggaggacttggccacacatgcccggg 300
          |||| ||||||  |||||||||||||| ||  ||||||||| || ||||| |||| |||
Db   5614 cgcccaccttggcctccccagtgctcc-agatatacggggct--gactac--cggc 5670

Qy     61 tctcaagctcaccgcaactcccgccctcccggggttcaagcgaattctcctggtctcagcctcc 120
          |||||| || |||  || |||| || |||||||||||||||||| ||||||||||||||||
Db   5380 ttttttttctgagatagg-tttcaactcttgttgcccaggctgggtggtcaatggcgcaa 5438

Qy    301 tggctcaattttgtgttgtaagaaacaggggtttcactgatgtgcccaagctggtctcctg-- 357
          ||||||||||| ||||||| | |  |||||| || ||||||||||| | ||||||
Db   5439 tcttggctcactgcaacctctcgccttccgggttcaagtgattctctgcctcagcctcc 5498

Qy    358 --agctcaaagcagtcacactgtcctcagctcagcctcccaaagtgctggggattacaggcgtgcagc 415
          |||||||||||||| || |||  || | |||||||| |||| | |||||||||||
Db   5671 tttttttttttgttct-gagacacagttcactctgttacccaggctggagtagggtgg 5729

Qy    121 cagtagctgggattacaggcatgtgccacccacgctcggctaatttttgtatttttttttag 180
          |||| ||||||||||| |||||| |||||||||||| | ||||||||||||||||||
Db   5439 aagtaactgggattacaggaggaagtgcaccaacccagctaattt---tgtatttttag 5555

Qy    181 tagagatggagttctccatgtgttggtcaggctggtctcgaactcccagatc 240
          ||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db   5499 aagtaactgggattacaggaggaagtgcaccaacccagctaattt---tgtatttttag 5555

Qy    476 tgaagtgcaagtgtgtgatca--caagctcactgcagccctcaaactgcggaagtcaag-ca 532
          ||||| | ||||||| || ||| |||||||||||||||  || ||||||||||| |||| ||
Db   5790 ctcccaagtagccgagattacaggcatgtgccacccacaccaagtaattttgtatttt 5849

Qy    416 cgtcgcctggccttttttattttttttaagacaacaggtgtccacactcttacccagga 475
          ||| | |||| ||||| ||||||||||||||||||||| ||| ||||| |||||||
Db   5730 cctgatctcggatcaactgcaacctccgcctctgggctcaagtgattgcctgttcagc 5789

Qy    533 tcctcctgcctcagcctcccaagtagctgagactggggacccaaagacaatgcacacactacaccgt 591
          ||| ||||| |||||||| |||| ||||||||| ||||| |||||||| ||| ||||||||
Db   5910 tccaccgcctcagcctcccaaagtgctgagattataggtgtgagccaccacgggcc 5969

Qy    592 ------------taatttttattttttattttattatttttgagacagagtctcaactctgt 640
          |||||||||||||||||||||||||||||| ||||| || ||||| |||| ||
Db   6028

Qy    641 caccccaggctggagtgcagtggcgcgaatcttggctcactgcaacctctgcctcccggtt 700
          | ||||||| ||||||||| ||||| ||| || ||| ||| || ||||||||||||||| |||
Db   5970 tcaggaagtgtattttattttttaaattaattattattgagatggagtct-tgctctgt 6028

Qy    701 caagttattctcctgccccaggctcctgagtagctggactacaggcgc---ccaccacg 757
          |||| ||||||||||| |||||||||||||| | |||||||| || |||||||
Db   6089 caagccattcctgcctcagcctcccagagtagctgggactacaggcgccccaccacca 6148

Qy    758 cctagctaattttttgtatttttagtagagatggg-gttcaccaatgttcgccaggttga 816
          |||| ||||||||||||| |||||||||||||| | |||||| || |||| ||||| || |
Db   6149 ccggtcaattttttgtattttttagtagagacgggg-ttttcaccgtgttagccaggatggg 6208

Qy    817 tcttgatctctggacctgtgatctgcctcgcctcccgaaagtgctgggattacag 876
          ||||||||| ||||| ||| ||| ||||| ||||||||| |||||||||||| |||
Db   6209 tcttgatctctgacctcgtgatctcggcctcccaaagtgctgggattacag 6268
```

us-09-964-678a-1.rng

```
Qy    877 gcgtgagccaccacgcccggcttattttaattttgtttgttgaaatggaatcttcact 936
          | ||||||||||||||||| |||||||
Db   6269 gtgtgagccaccaccacccggct------atttttattttttgagacagggactcact 6320

Qy    937 ctgttacccaggctggagtgcaatgccaaatctcggctcaactctgcctccg 996
          ||||| |||    ||||||||||||||| ||||  ||  |   |||||| ||||||
Db   6321 ctgtcacctgggctgcagtgcagtgg-tacaccatagctcactgcagcctcgaactcctg 6379

Qy    997 ggctcaagcgattcctcctgtctcagcctcccaagcagctgggattacgggcacctgccac 1056
          ||||||| |||| ||| ||||  |||||| ||||| |||||| |||||| |||||
Db   6380 agctcaagtgatcctcccacctcagcctcccacaagtaattgggactacaggtgcacccac 6439

Qy   1057 cacaccccgctaatttttgtat-----------tttcattagagacgggggtttcac 1101
          || ||| ||||||| | |||           |||  ||||| -||-||||
Db   6440 catgcccacctaatttattattttattttcatagagatgagggttccc 6499

Qy   1102 catatttgtcaggctggtctcaaactcctgacctcaggtgacccacctgcctcagccttc 1161
          | ||| ||||||||||| |||||||| | |||| ||||| |||| |
Db   6440 catgcccacctaatttattattttattttcatagagatgagggttccc 6499

Qy   1162 caaagtgctgggattacaggcgtgagccacctcaccagccggctaatttagataaa 1218
          ||||||||||| |||||||| |||||||| ||||| | ||||
Db   6560 caaagtgctgagattacaggcatgagccaccgtgcccagctaggaatcattttaa 6616
```

Search completed: April 24, 2002, 19:42:15
Job time: 3739 sec